

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:13:56 ; Search time 137 Seconds

(without alignments)

5197.647 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITASQQTGRLGCIIT.....PMWCLLLSVGVGIIYLLPNR 1985

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq\_23Sep04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	10465	100.0	1985	5	AAO18001	Aao18001 Hepatitis
2	10465	100.0	1985	5	AAE15717	Aae15717 Hepatitis
3	10465	100.0	1985	8	ADJ57846	Adj57846 HCV repli
4	10462	100.0	1985	5	AAE15729	Aae15729 Hepatitis
5	10462	100.0	2201	5	ABG30581	Abg30581 Hepatitis
6	10462	100.0	2201	5	ABG30587	Abg30587 Hepatitis
7	10462	100.0	3010	5	ABG32451	Abg32451 Hepatitis
8	10460	100.0	1985	5	AAE15731	Aae15731 Hepatitis
9	10460	100.0	1985	5	AAE15730	Aae15730 Hepatitis
10	10459	99.9	1985	5	AAE15730	Aae15730 Hepatitis
11	10458	99.9	1985	5	AAE15727	Aae15727 Hepatitis
12	10458	99.9	3010	5	ABG32460	Abg32460 Hepatitis
13	10458	99.9	3010	5	ABG32454	Abg32454 Hepatitis
14	10457	99.9	3010	5	ABG32461	Abg32461 Hepatitis
15	10456	99.9	3010	5	ABG32458	Abg32458 Hepatitis
16	10456	99.9	3010	5	ABG32459	Abg32459 Hepatitis
17	10456	99.9	3010	5	ABG32455	Abg32455 Hepatitis
18	10456	99.9	3010	5	ABG32457	Abg32457 Hepatitis
19	10456	99.9	3010	5	ABG32452	Abg32452 Hepatitis
20	10455	99.9	1985	5	AAE15720	Aae15720 Hepatitis
21	10455	99.9	3010	5	ABG32453	Abg32453 Hepatitis
22	10451.5	99.9	3011	5	ABG32456	Abg32456 Hepatitis
23	10450	99.9	1985	5	AAE15722	Aae15722 Hepatitis
24	10447	99.8	2201	5	ABG30582	Abg30582 Hepatitis
25	10442	99.8	2201	5	ABG30580	Abg30580 Hepatitis

#### RESULT 1

AAO18001

ID AAO18001 standard; protein; 1985 AA.

XX AC AAO18001;

XX DT 30-AUG-2002 (first entry)

XX DE Hepatitis C virus NS3 proteinase/helicase.

XX KW Hepatitis; HCV; core-neo; NS3 proteinase/helicase; diagnosis;

XX KX virucide; hepatotropic; gene therapy; anti-viral; enzyme.

XX OS Hepatitis C virus.

XX PN WO200238793-A2.

XX PD 16-MAY-2002.

XX PF 02-NOV-2001; 2001WO-US046350.

XX PR 07-NOV-2000; 2000US-0245866P.

XX PA (ANAD-) ANADYS PHARM INC.

XX PI Bichko V;

XX DR WPI; 2002-490082/52.

XX N-PSDB; AAL47276.

XX FT Novel nucleic acid encoding replication competent recombinant hepatitis C

XX FT virus genome useful for screening anti-hepatitis C virus therapeutics and

XX PS Claim 6; Page 50-51; 85pp; English.

XX CC The present invention provides protein and coding sequences from

XX CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and

XX CC able to replicate efficiently when transfected into a susceptible cell

XX CC line without reducing the growth rate of the cell line by more than 10

XX CC fold. The sequences are useful for screening for anti-HCV therapeutics,

XX CC for detecting antibodies to HCV in a biological sample such as blood,

XX CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,

XX CC for deriving authentic HCV components such as replication-component non-

XX CC infectious, replication-defective infection-component, and replication-

XX CC defective non-infectious HCV, in gene therapy or gene vaccination

XX CC targeted to hepatic tissue for treating an animal infected or susceptible

XX CC to HCV infection and for studying HCV infection and propagation. The

#### ALIGNMENTS

26	10441	99.8	2201	5	ABG30594	Hepatitis
27	10439	99.8	2201	5	ABG30584	Hepatitis
28	10439	99.8	2201	5	ABG30589	Hepatitis
29	10438	99.7	2201	5	ABG30591	Hepatitis
30	10438	99.7	2201	5	ABG30593	Hepatitis
31	10438	99.7	2201	5	ABG30590	Hepatitis
32	10437	99.7	2201	5	ABG30598	Hepatitis
33	10437	99.7	2201	5	ABG30595	Hepatitis
34	10436	99.7	2201	5	ABG30601	Hepatitis
35	10435	99.7	2201	5	ABG30602	Hepatitis
36	10435	99.7	2201	5	ABG30599	Hepatitis
37	10434	99.7	2201	5	ABG30585	Hepatitis
38	10434	99.7	2201	5	ABG30600	Hepatitis
39	10428	99.6	2201	5	ABG30583	Hepatitis
40	10421	99.6	2201	5	ABG30586	Hepatitis
41	10421	99.6	2201	5	ABG30588	Hepatitis
42	10392	99.3	2063	7	ADD67963	Hepatitis
43	10202	97.5	1985	6	ABU09575	HCV Met-N
44	10181	97.3	1985	6	ABU09574	HCV Met-N
45	10170	97.2	2201	2	AAW01680	HCV NS2-N

CC present sequence is the HCV NS3 proteinase/helicase

XX Sequence 1985 AA;

Query Match 100.0%; Score 10465; DB 5; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	M	A	P	I	T	A	S	Q	O	T	R	G	L	L	G	C	I	T	S	L	T	G	R	N	Q	V	E	G	V	Q	V	S	T	A	T	Q	S	F	L	A	T	C	V	N	G	V	C	W	T	V	H	G	A	60						
DB	1	M	A	P	I	T	A	S	Q	O	T	R	G	L	L	G	C	I	T	S	L	T	G	R	N	Q	V	E	G	V	Q	V	S	T	A	T	Q	S	F	L	A	T	C	V	N	G	V	C	W	T	V	H	G	A	60						
QY	61	G	S	T	L	A	G	P	K	P	I	T	M	N	V	D	D	L	V	G	M	A	P	P	G	A	S	L	T	P	C	T	G	S	S	D	L	V	T	R	H	A	D	V	I	P	V	R	R	120											
DB	61	G	S	T	L	A	G	P	K	P	I	T	M	N	V	D	D	L	V	G	M	A	P	P	G	A	S	L	T	P	C	T	G	S	S	D	L	V	T	R	H	A	D	V	I	P	V	R	R	120											
QY	121	G	D	S	R	G	S	L	L	P	R	P	V	S	Y	L	K	G	S	S	G	P	L	L	C	P	S	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	V	D	P	V	P	S	E	M	T	M	180					
DB	121	G	D	S	R	G	S	L	L	P	R	P	V	S	Y	L	K	G	S	S	G	P	L	L	C	P	S	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	V	D	P	V	P	S	E	M	T	M	180					
QY	181	R	S	P	V	F	T	D	N	S	P	P	A	V	P	O	T	F	O	V	A	H	L	H	A	P	T	G	S	K	T	K	V	P	A	A	Y	A	A	Q	K	V	L	N	P	S	V	A	A	T	L	G	F	240							
DB	181	R	S	P	V	F	T	D	N	S	P	P	A	V	P	O	T	F	O	V	A	H	L	H	A	P	T	G	S	K	T	K	V	P	A	A	Y	A	A	Q	K	V	L	N	P	S	V	A	A	T	L	G	F	240							
QY	241	A	N	M	S	K	A	G	I	D	P	N	I	R	T	G	V	R	T	T	I	T	C	A	P	I	T	S	T	Y	G	K	F	L	A	D	G	G	S	G	A	Y	D	I	I	C	D	E	C	H	S	T	D	S	T	300					
DB	241	A	N	M	S	K	A	G	I	D	P	N	I	R	T	G	V	R	T	T	I	T	C	A	P	I	T	S	T	Y	G	K	F	L	A	D	G	G	S	G	A	Y	D	I	I	C	D	E	C	H	S	T	D	S	T	300					
QY	301	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	A	T	P	G	S	V	T	V	P	H	P	N	I	E	E	V	A	L	S	T	G	E	I	P	F	Y	K	A	I	P	I	E	T	360				
DB	301	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	A	T	P	G	S	V	T	V	P	H	P	N	I	E	E	V	A	L	S	T	G	E	I	P	F	Y	K	A	I	P	I	E	T	360				
QY	361	K	G	R	H	I	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	G	N	A	V	A	Y	R	G	L	D	V	S	I	P	T	S	G	D	V	I	V	A	T	A	L	M	T	G	F	420									
DB	361	K	G	R	H	I	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	G	N	A	V	A	Y	R	G	L	D	V	S	I	P	T	S	G	D	V	I	V	A	T	A	L	M	T	G	F	420									
QY	421	G	D	F	S	V	I	D	C	N	T	C	T	Q	T	V	T	F	S	L	D	T	F	T	I	T	T	V	P	Q	D	A	V	S	R	O	R	G	R	T	G	R	G	M	G	I	Y	R	F	V	T	P	G	480							
DB	421	G	D	F	S	V	I	D	C	N	T	C	T	Q	T	V	T	F	S	L	D	T	F	T	I	T	T	V	P	Q	D	A	V	S	R	O	R	G	R	T	G	R	G	M	G	I	Y	R	F	V	T	P	G	480							
QY	481	E	R	P	S	G	M	F	D	S	S	V	L	C	E	C	Y	D	A	G	A	C	A	M	Y	E	L	T	P	A	E	T	S	V	R	L	A	V	L	N	T	P	G	L	P	V	C	Q	D	H	L	E	F	540							
DB	481	E	R	P	S	G	M	F	D	S	S	V	L	C	E	C	Y	D	A	G	A	C	A	M	Y	E	L	T	P	A	E	T	S	V	R	L	A	V	L	N	T	P	G	L	P	V	C	Q	D	H	L	E	F	540							
QY	541	T	H	I	D	A	H	F	L	S	O	T	K	O	A	G	N	F	P	L	V	A	Y	O	A	T	V	C	A	R	A	Q	A	P	P	S	D	O	M	K	C	L	I	R	K	P	T	L	H	G	P	T	L	600							
DB	541	T	H	I	D	A	H	F	L	S	O	T	K	O	A	G	N	F	P	L	V	A	Y	O	A	T	V	C	A	R	A	Q	A	P	P	S	D	O	M	K	C	L	I	R	K	P	T	L	H	G	P	T	L	600							
QY	601	Y	R	L	G	A	V	Q	N	E	V	T	T	H	P	T	K	Y	I	M	A	C	M	S	A	D	L	E	V	T	T	S	T	V	L	V	G	V	L	A	A	A	Y	C	L	T	T	G	S	V	I	V	G	660							
DB	601	Y	R	L	G	A	V	Q	N	E	V	T	T	H	P	T	K	Y	I	M	A	C	M	S	A	D	L	E	V	T	T	S	T	V	L	V	G	V	L	A	A	A	Y	C	L	T	T	G	S	V	I	V	G	660							
QY	661	I	I	L	S	G	P	A	I	P	D	R	E	V	L	Y	R	E	P	D	E	M	E	C	A	S	H	L	P	Y	I	E	Q	G	M	L	A	E	O	K	Q	K	A	I	G	L	L	O	T	A	T	K	O	A	E	720					
DB	661	I	I	L	S	G	P	A	I	P	D	R	E	V	L	Y	R	E	P	D	E	M	E	C	A	S	H	L	P	Y	I	E	Q	G	M	L	A	E	O	K	Q	K	A	I	G	L	L	O	T	A	T	K	O	A	E	720					
QY	721	A	A	P	V	S	K	M	T	L	E	A	F	A	K	M	N	N	I	S	G	I	O	Y	L	A	G	I	S	T	P	G	N	P	A	I	A	S	L	M	A	F	T	A	S	I	T	S	P	L	T	O	H	780							
DB	721	A	A	P	V	S	K	M	T	L	E	A	F	A	K	M	N	N	I	S	G	I	O	Y	L	A	G	I	S	T	P	G	N	P	A	I	A	S	L	M	A	F	T	A	S	I	T	S	P	L	T	O	H	780							
QY	781	T	L	L	F	N	I	L	G	V	A	Q	L	A	P	P	S	A	S	A	F	V	G	A	I	A	G	A	V	G	S	I	G	L	K	V	L	V	D	I	L	A	G	Y	C	A	G	V	A	G	A	L	V	A	840						
DB	781	T	L	L	F	N	I	L	G	V	A	Q	L	A	P	P	S	A	S	A	F	V	G	A	I	A	G	A	V	G	S	I	G	L	K	V	L	V	D	I	L	A	G	Y	C	A	G	V	A	G	A	L	V	A	840						
QY	841	F	K	M	S	G	E	M	P	S	T	E	D	L	V	N	L	P	A	I	L	S	P	G	A	L	V	G	V	W	C	A	A	I	L	R	H	V	G	P	O	E	G	A	V	Q	M	N	R	L	I	A	P	A	S	900					
DB	841	F	K	M	S	G	E	M	P	S	T	E	D	L	V	N	L	P	A	I	L	S	P	G	A	L	V	G	V	W	C	A	A	I	L	R	H	V	G	P	O	E	G	A	V	Q	M	N	R	L	I	A	P	A	S	900					
QY	901	G	N	H	V	S	P	H	Y	P	E	S	D	A	A	R	V	T	Q	I	L	S	S	L	T	I	T	Q	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	M	L	R	D	V	M	D	W	I	C	960			
DB	901	G	N	H	V	S	P	H	Y	P	E	S	D	A	A	R	V	T	Q	I	L	S	S	L	T	I	T	Q	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	M	L	R	D	V	M	D	W	I	C	960			
QY	961	T	V	L	T	D	F	K	T	W	L	Q	S	K	L	P	R	L	P	G	V	P	P	F	S	C	O	R	G	Y	K	G	V	M	R	G	D	G	I	M	Q	T	T	C	P	C	G	A	Q	I	T	G	H	V	K	N	G	S	M	R	1020
DB	961	T	V	L	T	D	F	K	T	W	L	Q	S	K	L	P	R	L	P	G	V	P	P	F	S	C	O	R	G	Y	K	G	V	M	R	G	D	G	I	M	Q	T	T	C	P	C	G	A	Q	I	T	G	H	V	K	N	G	S	M	R	1



AAE15717  
 ID AAE15717 standard; protein; 1985 AA.  
 XX  
 AC AAE15717;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Hepatitis C virus (HCV) replBartMan polyprotein.  
 XX  
 KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
 KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
 KW replBartMan.  
 XX  
 OS Hepatitis C virus.  
 XX  
 XX WO200189364-A2.  
 PN  
 XX 29-NOV-2001.  
 PD  
 XX 23-MAY-2001; 2001WO-US016822.  
 PF  
 XX 23-MAY-2000; 2000US-00576989.  
 PR  
 XX (UNIW ) UNIV WASHINGTON.  
 PA  
 XX Rice CM, Blight KJ;  
 PI  
 XX WPI; 2002-066755/09.  
 DR  
 XX N-PSDB; AAD25322.  
 DR  
 XX  
 XX Hepatitis C virus variants having greater transfection efficiency and  
 PT ability to survive subpassage, useful as a vaccine for immunizing primate  
 PT to the virus, comprise non-naturally occurring viral sequences.  
 XX  
 XX Claim 14; Page 65; 174pp; English.  
 XX  
 CC The invention relates to Hepatitis C virus (HCV) variants which include  
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
 CC variants that have a transfection efficiency and ability to survive  
 CC subpassage greater than HCV that have wild-type polyprotein coding  
 CC regions. The polynucleotides of the invention are useful for identifying  
 CC a cell line that is permissive for infection with HCV and detecting  
 CC replication of HCV in cells of the cell line. They are also useful for  
 CC testing a compound for anti-viral properties and for inhibiting HCV  
 CC infection. They are also useful for the generation of defined HCV virus  
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
 CC attachment, penetration and entry, structure/function studies on HCV  
 CC proteins and RNA elements and identification of new antiviral targets, a  
 CC systematic survey of cell culture systems and conditions to identify  
 CC those that support wild-type and variant HCV RNA replication and particle  
 CC release, production of adaptive HCV variants capable of more efficient  
 CC replication in cell culture, production of HCV variants with altered  
 CC tissue or species tropism, establishment of alternative animal models for  
 CC inhibitor evaluation including those supporting HCV variant replication,  
 CC development of cell-free HCV replication assays, production of  
 CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
 CC derivatives as possible vaccine candidates, engineering of attenuated or  
 CC defective HCV derivatives for expression of heterologous gene products  
 CC for gene therapy and vaccine applications and for utilisation of the HCV  
 CC glycoproteins for targeted delivery of therapeutic agents to the liver  
 CC or other cell types with appropriate receptors. Vaccine comprising these  
 CC sequences is useful for inducing immunoprotection to HCV in a primate.  
 CC The present sequence is Hepatitis C virus (HCV) replBartMan polyprotein  
 XX  
 XX Sequence 1985 AA;  
 SQ  
 Query Match 100.0%; Score 10465; DB 5; Length 1985;  
 Best Local Similarity 100.0%; Pred No 0;  
 Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGVQVVSSTATQSFATCNGVCWTVYHGA 60  
 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGVQVVSSTATQSFATCNGVCWTVYHGA 60

QY	61	GSKTLAGPKG	ITQMYTNVDQDLVGMQAPPGARSLT	PCTCGSSDLXLVTRHADVPVRR	120
Db	61	GSKTLAGPKG	ITQMYTNVDQDLVGMQAPPGARSLT	PCTCGSSDLXLVTRHADVPVRR	120
QY	121	GDSEGLLSRP	SVYLGSSGRLCPGSHAVGIFRAAVCT	RGVAKAVDFVFSMETTM	180
Db	121	GDSEGLLSRP	SVYLGSSGRLCPGSHAVGIFRAAVCT	RGVAKAVDFVFSMETTM	180
QY	181	RSPVFTD	NSPPAVPQTFQVAHLHAPTGS	GSKSTKVPAAAYAAQGYKVLVNSVAATLGF	240
Db	181	RSPVFTD	NSPPAVPQTFQVAHLHAPTGS	GSKSTKVPAAAYAAQGYKVLVNSVAATLGF	240
QY	241	AYMSKAH	GIDPNIRKTGVRTITTGAPIT	YSTYVKFLADGCGSGGAYDIIICDECHSTDT	300
Db	241	AYMSKAH	GIDPNIRKTGVRTITTGAPIT	YSTYVKFLADGCGSGGAYDIIICDECHSTDT	300
QY	301	ILGIGTVL	DOAETAGARLVLATATPPGS	VTVPHPNIEEVALSSTGIBPFYKAIPIETI	360
Db	301	ILGIGTVL	DOAETAGARLVLATATPPGS	VTVPHPNIEEVALSSTGIBPFYKAIPIETI	360
QY	361	KGGRHLIF	CHSKKKCDELAALKSLGLN	VAYYRGLDVSVIPTSGDVIVVATDALMTGFT	420
Db	361	KGGRHLIF	CHSKKKCDELAALKSLGLN	VAYYRGLDVSVIPTSGDVIVVATDALMTGFT	420
QY	421	GDSDVID	CNTCVTQTVD	FSLDPTFTTITTTVPQDAVSR	480
Db	421	GDSDVID	CNTCVTQTVD	FSLDPTFTTITTTVPQDAVSR	480
QY	481	ERPSG	WFDSSVLCEDYDAGCANYEL	TPAETSRLRAYLNTPLPVCODHLEFWESVFTGL	540
Db	481	ERPSG	WFDSSVLCEDYDAGCANYEL	TPAETSRLRAYLNTPLPVCODHLEFWESVFTGL	540
QY	541	THIDA	FLPSOTKAGDNFPYLVAY	QATVCARAQAPPSWDMKCLRLKPLTGHSP	600
Db	541	THIDA	FLPSOTKAGDNFPYLVAY	QATVCARAQAPPSWDMKCLRLKPLTGHSP	600
QY	601	YRLGAV	QVNEVTTTHPIKTI	MACMSADLEVTSTWLVGGVLAALAAAYCLT	660
Db	601	YRLGAV	QVNEVTTTHPIKTI	MACMSADLEVTSTWLVGGVLAALAAAYCLT	660
QY	661	IIISG	KEAIPDREVL	YREFDEMEBCASHLPYIEQGMQLAEQKQK	720
Db	661	IIISG	KEAIPDREVL	YREFDEMEBCASHLPYIEQGMQLAEQKQK	720
QY	721	AAPV	ESKWRTL	LEAFWAKHWNFI	840
Db	721	AAPV	ESKWRTL	LEAFWAKHWNFI	840
QY	781	TLFNL	IGWVAQAQLAPP	SAASAFVAGIAGAAVSGIGLKVLDILAGY	840
Db	781	TLFNL	IGWVAQAQLAPP	SAASAFVAGIAGAAVSGIGLKVLDILAGY	840
QY	841	FKVMS	GEMPTEDLVNLLPAIL	SPGALVGVVCAATLRRHVGE	900
Db	841	FKVMS	GEMPTEDLVNLLPAIL	SPGALVGVVCAATLRRHVGE	900
QY	901	GNHVS	THVPVPSDAAARVTQ	ILSSLTITQLKRLHWNEDCSTPCSSGLRDV	960
Db	901	GNHVS	THVPVPSDAAARVTQ	ILSSLTITQLKRLHWNEDCSTPCSSGLRDV	960
QY	961	TVLTD	FKTLQSKLLPRL	PGVPPFFSCQGYKGVWRGDGIMQITTC	1020
Db	961	TVLTD	FKTLQSKLLPRL	PGVPPFFSCQGYKGVWRGDGIMQITTC	1020
QY	1021	IVGPR	CSNTHGTPPINAYTT	GPCTPSPAPNSRALMRVAEEYVEVTR	1080
Db	1021	IVGPR	CSNTHGTPPINAYTT	GPCTPSPAPNSRALMRVAEEYVEVTR	1080
QY	1081	TTDNV	KCPQVPAPEFTEVD	GVRLHRYAPACKELLREEVTVLGLNOVL	1140
Db	1081	TTDNV	KCPQVPAPEFTEVD	GVRLHRYAPACKELLREEVTVLGLNOVL	1140

1141 PDVAULTSMLTDPSPHITAKRRLARGSPPLSSASSQSAPSIKATCTTRHDSPAD 1200  
 1141 PDVAULTSMLTDPSPHITAKRRLARGSPPLSSASSQSAPSIKATCTTRHDSPAD 1200  
 1201 LLEANLLWRQEMGNITRVESNKVVILDSFPLQAEEDEREVSVPALILRSRKFPFRAM 1260  
 1201 LLEANLLWRQEMGNITRVESNKVVILDSFPLQAEEDEREVSVPALILRSRKFPFRAM 1260  
 1261 PIWARPDPNPPLESWKPDYVPPVHGCPLPAPAKAPP:PPPRKRRTVVLSSTVSSALA 1320  
 1261 PIWARPDPNPPLESWKPDYVPPVHGCPLPAPAKAPP:PPPRKRRTVVLSSTVSSALA 1320  
 1321 ELATKTFGSSSADVSGTATASPOPSDDGAGSDVESYSSMPPLGEGPDPDSDGSW 1380  
 1321 ELATKTFGSSSADVSGTATASPOPSDDGAGSDVESYSSMPPLGEGPDPDSDGSW 1380  
 1381 STVSEASDVVCCGSYSYTWTCALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
 1381 STVSEASDVVCCGSYSYTWTCALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
 1441 LRQKKVTFORLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
 1441 LRQKKVTFORLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
 1501 DVNLSKAVNHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLG 1560  
 1501 DVNLSKAVNHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLG 1560  
 1561 VRVCEKALYDVVSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 1561 VRVCEKALYDVVSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 1621 TVTENDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRAGSVLT 1680  
 1621 TVTENDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRAGSVLT 1680  
 1681 TSCGNTLTCYLKAAACRAKLODCTMLVCGDDLVCESAGTQDEASLRAPTEAMTRY 1740  
 1681 TSCGNTLTCYLKAAACRAKLODCTMLVCGDDLVCESAGTQDEASLRAPTEAMTRY 1740  
 1741 SAPPGPDPKPEYDLELITSCSSNVSVVAHDASGRVYVLTDRDPTTPLARAWEATARHTPVN 1800  
 1741 SAPPGPDPKPEYDLELITSCSSNVSVVAHDASGRVYVLTDRDPTTPLARAWEATARHTPVN 1800  
 1801 SWLGNTIMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPIQIOL 1860  
 1801 SWLGNTIMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPIQIOL 1860  
 1861 HGLSFLSHSYSGEINRVASCLRLGVPLRVWRHRARSVRARLLSQGGRAATCGKLYF 1920  
 1861 HGLSFLSHSYSGEINRVASCLRLGVPLRVWRHRARSVRARLLSQGGRAATCGKLYF 1920  
 1921 NNAVTRKLLTPIPAASQDLSSWVFAGYSGGDIYHLSLRARPRWFMWCLLLLSVGVIY 1980  
 1921 NNAVTRKLLTPIPAASQDLSSWVFAGYSGGDIYHLSLRARPRWFMWCLLLLSVGVIY 1980  
 1981 LLPNR 1985  
 1981 LLPNR 1985

RESULT 3  
 ADJ57846 standard; protein; 1985 AA.  
 ID ADJ57846;  
 AC ADJ57846;  
 XX ADJ57846;  
 XX ADJ57846;  
 DT 06-MAY-2004 (first entry)  
 XX HCV replicon.  
 DE hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide;  
 KW hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide;  
 KW HCV replicon.

XX Unidentified.  
 OS WO2004015131-A2.  
 XX 19-FEB-2004.  
 XX 12-AUG-2003; 2003WO-US025260.  
 XX 12-AUG-2002; 2002US-0402661P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX Gao M, Lemm JA, O'boyle DR, Nower P, Rigat K, Sun J;  
 PI WPI; 2004-180685/17.  
 PI N-PSDB; ADJ57845.  
 DR Use of hepatitis C virus assays or reporter assays, e.g. identifying a  
 DR compound that inhibits hepatitis C virus RNA replication or identifying a  
 DR compound that modulates the activity of a gene of interest.  
 XX Claim 4; SEQ ID NO 2; 45pp; English.  
 XX The present invention relates to the use of hepatitis C virus (HCV)  
 CC assays for identifying a compound that inhibits HCV RNA replication and  
 CC reporter assays for identifying a compound that modulates the activity of  
 CC a gene of interest. The assays are useful for identifying a compound that  
 CC inhibits HCV RNA replication or for identifying a compound that modulates  
 CC the activity of a gene of interest. The HCV assay is useful for high  
 CC throughput screening that quantifies both the amount of HCV RNA  
 CC replication inhibitory activity associated with a test compound and the  
 CC amount of cytotoxicity associated with the test compound. The compound is  
 CC useful for treating hepatitis C infection. Assays of the invention have  
 CC distinct advantages when compared to qRT-PCR or other methods in that  
 CC assays of the invention may take place in situ in a detergent based crude  
 CC cell lysate, which requires no further preparation prior to performing  
 CC the assays. The assays do not also involve numerous manipulations to add  
 CC or subtract reagents after addition of test compounds and are desirably  
 CC based on a viral protein which is required by the HCV replicon for  
 CC replication. The present sequence represents a HCV replicon used in the  
 CC assay of the invention.  
 XX Sequence 1985 AA;  
 SQ  
 Query Match 100.0%; Score 10465; DB 8; Length 1985;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPITAYSOQTRGLLCIITSLTGRDRNQVEGEVQVWVSTATQSFATCVNGVCMWTVYHGA 60  
 DB 1 MAPITAYSOQTRGLLCIITSLTGRDRNQVEGEVQVWVSTATQSFATCVNGVCMWTVYHGA 60  
 QY 61 GSKTLGPKGPIQMTYNTVDQDLVGMQAPPGARSLTPTCTCGSSDLXLVTRHADVIPVRR 120  
 DB 61 GSKTLGPKGPIQMTYNTVDQDLVGMQAPPGARSLTPTCTCGSSDLXLVTRHADVIPVRR 120  
 QY 121 GDSRGLLSRPVSYLKGSGGGLLCPSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180  
 DB 121 GDSRGLLSRPVSYLKGSGGGLLCPSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180  
 QY 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSKTKVPAAYAAQGYKVLVLPNSVAATLGF 240  
 DB 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSKTKVPAAYAAQGYKVLVLPNSVAATLGF 240  
 QY 241 AYMSKAGIDPNIRTVGRTITTTGAPITYSTYHKLADGGCGSGGAYDIIICDECHSTSTT 300  
 DB 241 AYMSKAGIDPNIRTVGRTITTTGAPITYSTYHKLADGGCGSGGAYDIIICDECHSTSTT 300  
 QY 301 ILGIGTVLDQAEATAGARLVVLTATATPGSVTVPHNIEEVALSGTGBIPFGKAIPIETI 360  
 DB 301 ILGIGTVLDQAEATAGARLVVLTATATPGSVTVPHNIEEVALSGTGBIPFGKAIPIETI 360

361 KGRHLIFCHSKKKDELAALSGLGNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Db LRLQKKVTFDRLQVLDHVRDVLKEMKAKASTVKAKLLSVEBACKLTPHRSARSKFYGAK 1500  
361 KGRHLIFCHSKKKDELAALSGLGNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Qy DVRLSSKAVNHIISVWKOLLEDTETPTDITIMAKNEVFCVQPEKGGKRPALRIVFPDLG 1560  
421 GDFSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRSQRRGRTGRGRGIYRFVTPG 480  
Db GDFSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRSQRRGRTGRGRGIYRFVTPG 480  
421 GDFSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRSQRRGRTGRGRGIYRFVTPG 480  
Qy ERPSGMFSSVLCYDAGCAWYELTPAETSVRURAYLNTGCLPVCQDHLFWFSVFTGL 540  
Db ERPSGMFSSVLCYDAGCAWYELTPAETSVRURAYLNTGCLPVCQDHLFWFSVFTGL 540  
481 ERPSGMFSSVLCYDAGCAWYELTPAETSVRURAYLNTGCLPVCQDHLFWFSVFTGL 540  
Qy THIDAHLSQTKQAGDNFPYLVAYQATVCARAQAPPSDWMKCLIRLKPTELHGPTLL 600  
Db THIDAHLSQTKQAGDNFPYLVAYQATVCARAQAPPSDWMKCLIRLKPTELHGPTLL 600  
541 THIDAHLSQTKQAGDNFPYLVAYQATVCARAQAPPSDWMKCLIRLKPTELHGPTLL 600  
Qy YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
Db YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
Qy IILSGKEPAILPDREVLREFDEMEECASHLPYIEQGMOLABOPKQKXGAILLOATKQAEA 720  
Db IILSGKEPAILPDREVLREFDEMEECASHLPYIEQGMOLABOPKQKXGAILLOATKQAEA 720  
721 AAPVVEKWTLEAFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db AAPVVEKWTLEAFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
781 TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
Db TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
841 FKVMSEMPSTEDLVNLLPAILSPGALVGVCAAIIILRRHVGEAGVQWNNRLIAPASR 900  
Qy FKVMSEMPSTEDLVNLLPAILSPGALVGVCAAIIILRRHVGEAGVQWNNRLIAPASR 900  
841 FKVMSEMPSTEDLVNLLPAILSPGALVGVCAAIIILRRHVGEAGVQWNNRLIAPASR 900  
Qy GNHVSPTHYVPESDAAARVQTILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
Db GNHVSPTHYVPESDAAARVQTILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
901 GNHVSPTHYVPESDAAARVQTILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
Qy TVLTDFTKTLQSKLLPRLPGVPFFSCQYKGVWVRGDMQTTCPGCAQITGHVKNXGSMR 1020  
Db TVLTDFTKTLQSKLLPRLPGVPFFSCQYKGVWVRGDMQTTCPGCAQITGHVKNXGSMR 1020  
1021 IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALRWVAAEYVEVTVRGDFHYVTGM 1080  
Qy IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALRWVAAEYVEVTVRGDFHYVTGM 1080  
1021 IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALRWVAAEYVEVTVRGDFHYVTGM 1080  
Qy TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPILLREEVTVLVLNQYLVSQLPCEPE 1140  
Db TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPILLREEVTVLVLNQYLVSQLPCEPE 1140  
1141 PDVAVLTSMLTDPHSHTAETAKERLARGSPSSASSASQLSAPSUKATCTTRHDSPPDAD 1200  
Qy PDVAVLTSMLTDPHSHTAETAKERLARGSPSSASSASQLSAPSUKATCTTRHDSPPDAD 1200  
1141 PDVAVLTSMLTDPHSHTAETAKERLARGSPSSASSASQLSAPSUKATCTTRHDSPPDAD 1200  
Qy LIEANLLMRQEMGNITRVESENKVVILDSFPLQAEDEDEREVSVPABILRRSRKFPFRAM 1260  
Db LIEANLLMRQEMGNITRVESENKVVILDSFPLQAEDEDEREVSVPABILRRSRKFPFRAM 1260  
1201 LIEANLLMRQEMGNITRVESENKVVILDSFPLQAEDEDEREVSVPABILRRSRKFPFRAM 1260  
Qy PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRRTVVLSESTVSSALA 1320  
Db PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRRTVVLSESTVSSALA 1320  
1261 PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRRTVVLSESTVSSALA 1320  
Qy ELATKTFGSSSESSAVDSGTATASPDQSDGAGSDVESYSMPPLEGEPDPLSDGSW 1380  
Db ELATKTFGSSSESSAVDSGTATASPDQSDGAGSDVESYSMPPLEGEPDPLSDGSW 1380  
1321 ELATKTFGSSSESSAVDSGTATASPDQSDGAGSDVESYSMPPLEGEPDPLSDGSW 1380  
Qy STVSEASEDVSVCSSMSYTWTCALITPCAAETKLPINALSNLRRHNLVYATTSRGAS 1440  
Db STVSEASEDVSVCSSMSYTWTCALITPCAAETKLPINALSNLRRHNLVYATTSRGAS 1440  
1381 STVSEASEDVSVCSSMSYTWTCALITPCAAETKLPINALSNLRRHNLVYATTSRGAS 1440  
Qy LRLQKKVTFDRLQVLDHVRDVLKEMKAKASTVKAKLLSVEACKLTPHRSARSKFYGAK 1500

Db LRLQKKVTFDRLQVLDHVRDVLKEMKAKASTVKAKLLSVEACKLTPHRSARSKFYGAK 1500  
Qy DVRLSSKAVNHIISVWKOLLEDTETPTDITIMAKNEVFCVQPEKGGKRPALRIVFPDLG 1560  
Db DVRLSSKAVNHIISVWKOLLEDTETPTDITIMAKNEVFCVQPEKGGKRPALRIVFPDLG 1560  
Qy VVRCCKMALYDVVSTLTPQAVMGSSYGFQSPGQRFVFLVNAWKAKCPMGFAYDTRCFDS 1620  
Db VVRCCKMALYDVVSTLTPQAVMGSSYGFQSPGQRFVFLVNAWKAKCPMGFAYDTRCFDS 1620  
1561 VVRCCKMALYDVVSTLTPQAVMGSSYGFQSPGQRFVFLVNAWKAKCPMGFAYDTRCFDS 1620  
Qy TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
Db TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
Qy TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLAVICESAGTOEDEASLRAFTTEAMTRY 1740  
Db TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLAVICESAGTOEDEASLRAFTTEAMTRY 1740  
1681 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLAVICESAGTOEDEASLRAFTTEAMTRY 1740  
Qy SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGKRVVYLTTRDPTTPLARAAMETARHTPVN 1800  
Db SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGKRVVYLTTRDPTTPLARAAMETARHTPVN 1800  
1741 SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGKRVVYLTTRDPTTPLARAAMETARHTPVN 1800  
Qy SWLGNIIIMYAPTWARMLMTHFFSILLAOEQLKALDQCIYGACYSIEPLDLPQIIORL 1860  
Db SWLGNIIIMYAPTWARMLMTHFFSILLAOEQLKALDQCIYGACYSIEPLDLPQIIORL 1860  
1801 SWLGNIIIMYAPTWARMLMTHFFSILLAOEQLKALDQCIYGACYSIEPLDLPQIIORL 1860  
Qy HGLSAFSLHSYSFGEINRVASCLRKLGVPPLRVWRHRSVRARLLSOGGAAATCGKYL 1920  
Db HGLSAFSLHSYSFGEINRVASCLRKLGVPPLRVWRHRSVRARLLSOGGAAATCGKYL 1920  
1861 HGLSAFSLHSYSFGEINRVASCLRKLGVPPLRVWRHRSVRARLLSOGGAAATCGKYL 1920  
Qy NWAVRTKLTPIPAASQLDSSFWAGYSGGDIYHSLSRPRWFWMCILLLSVGVGIY 1980  
Db NWAVRTKLTPIPAASQLDSSFWAGYSGGDIYHSLSRPRWFWMCILLLSVGVGIY 1980  
1921 NWAVRTKLTPIPAASQLDSSFWAGYSGGDIYHSLSRPRWFWMCILLLSVGVGIY 1980  
Qy LIPNR 1985  
Db LIPNR 1985  
1981 LIPNR 1985  
RESULT 4  
AAE15729  
ID AAE15729 standard; protein; 1985 AA.  
XX AAE15729;  
AC AC  
XX XX  
DT 12-MAR-2002 (first entry)  
XX XX  
DE Hepatitis C virus (HCV) replbBartMan polyprotein variant A1174S.  
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KW replbBartMan; mutant; mutein; variant.  
XX XX  
OS Hepatitis C virus.  
OS Synthetic.  
XX XX  
FH Key Location/Qualifiers  
FT Misc-difference 1174  
ET /note= "Wild type Ala substituted with Ser"  
XX XX  
PN WO200189364-A2.  
XX XX  
PD 29-NOV-2001.  
XX XX  
PF 23-MAY-2001; 2001WO-US016822.  
XX XX  
PR 23-MAY-2000; 2000US-00576989.  
XX XX  
PA (UNIW ) UNIV WASHINGTON.  
XX XX  
PI Rice CM, Blight KJ;  
XX XX

DR WPI; 2002-066755/09.  
XX Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
PS Claim 14; Page; 174pp; English.  
XX  
XX The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of heterologous gene products  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) replicon/Man polyprotein  
CC variant. Note: The present sequence is not shown in the specification but  
CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the  
CC specification  
XX  
SQ Sequence 1985 AA;

Query Match 100.0%; Score 10462; DB 5; Length 1985;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIIITSLTGRDRNOVEGEVQVVSSTATQSFATCVNGVCWTYYHGA 60  
DB 1 MAPITAYSQOTRGLLCIIITSLTGRDRNOVEGEVQVVSSTATQSFATCVNGVCWTYYHGA 60  
QY 61 GSKTLAGPKGPTITQMTYNDQDLVGHQAPPGARSLTFCCTGSSDLYLVRHADVIPVRR 120  
DB 61 GSKTLAGPKGPTITQMTYNDQDLVGHQAPPGARSLTFCCTGSSDLYLVRHADVIPVRR 120  
QY 121 GDSRGLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
DB 121 GDSRGLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQYKVLVLPNSVAATLGF 240  
DB 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQYKVLVLPNSVAATLGF 240  
QY 241 AYSKAGIDPNRTGVRTITTCAPTITYTYGKFLADGGCGGAYDIIICDECHSTDSST 300  
DB 241 AYSKAGIDPNRTGVRTITTCAPTITYTYGKFLADGGCGGAYDIIICDECHSTDSST 300  
QY 301 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNIEEVALSGTBIPFYGAIPETI 360  
DB 301 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNIEEVALSGTBIPFYGAIPETI 360  
QY 361 KGGSHLLIFCHSKKCKDELAALSGLGNNAVAYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
DB 361 KGGSHLLIFCHSKKCKDELAALSGLGNNAVAYRGLDVSVIPTSGDVIVVATDALMTGFT 420

QY 421 GDFSDVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRSRORRGTGRNGIYRFTVTPG 480  
DB 421 GDFSDVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRSRORRGTGRNGIYRFTVTPG 480  
QY 481 ERPSGMDFSSVLCBVCYDAGCAWYELTTPAETSRLRAYLNTPLPVCODHLEFWSVFTGL 540  
DB 481 ERPSGMDFSSVLCBVCYDAGCAWYELTTPAETSRLRAYLNTPLPVCODHLEFWSVFTGL 540  
QY 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDQWKKLIRLKLKTLHGPTLL 600  
DB 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDQWKKLIRLKLKTLHGPTLL 600  
QY 601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTGSVIVGR 660  
DB 601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTGSVIVGR 660  
QY 661 IILSGKPAIIPDREVLVREFDEMEBECASHLYPIBQGMQLABQFKOKAIGLQATKQAEA 720  
DB 661 IILSGKPAIIPDREVLVREFDEMEBECASHLYPIBQGMQLABQFKOKAIGLQATKQAEA 720  
QY 721 AAPVVEKWRITLPAFWAKHMNFISGQIYLAGLSTLPGNPAIASLMAFTASITSLTQGH 780  
DB 721 AAPVVEKWRITLPAFWAKHMNFISGQIYLAGLSTLPGNPAIASLMAFTASITSLTQGH 780  
QY 781 TILFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
DB 781 TILFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
QY 841 FKVMGEMSPSTEDLVNLLPAILSPCALVGVVCAILRHHVGPGEVQVMMNRLTAFASR 900  
DB 841 FKVMGEMSPSTEDLVNLLPAILSPCALVGVVCAILRHHVGPGEVQVMMNRLTAFASR 900  
QY 901 GNHVSPTHVPESDAAAVTQLSLTITQLLKRHLQWINEDECSTFCSSGSLRDVMDTFC 960  
DB 901 GNHVSPTHVPESDAAAVTQLSLTITQLLKRHLQWINEDECSTFCSSGSLRDVMDTFC 960  
QY 961 TVLTDFKTLQSKLPLPLPGVFFPCQRYKGVWRGDGIMQTTCPGQAITGHVKNKSMR 1020  
DB 961 TVLTDFKTLQSKLPLPLPGVFFPCQRYKGVWRGDGIMQTTCPGQAITGHVKNKSMR 1020  
QY 1021 IVGPRTCSTNTHGTPPINAYTTGPTCTSPAPNYSRALWRAAEVVEVTRVGDVPHVTGM 1080  
DB 1021 IVGPRTCSTNTHGTPPINAYTTGPTCTSPAPNYSRALWRAAEVVEVTRVGDVPHVTGM 1080  
QY 1081 TTDNWKPCQVPAPEFTEVDGVRLHRYAPACKPLLREBVTFLVGLNQYLVGSQLPCEPE 1140  
DB 1081 TTDNWKPCQVPAPEFTEVDGVRLHRYAPACKPLLREBVTFLVGLNQYLVGSQLPCEPE 1140  
QY 1141 PDVAVLTSLMTDPSHITATKRRRLARGSPPSLASSASQLSAPSLKATCTTRHDSFDAD 1200  
DB 1141 PDVAVLTSLMTDPSHITATKRRRLARGSPPSLASSASQLSAPSLKATCTTRHDSFDAD 1200  
QY 1201 LIEANLLWQEMGNTTRVSEKVVILDSFEPLQAEDEREVSVPAEILRRSRKFPKRAM 1260  
DB 1201 LIEANLLWQEMGNTTRVSEKVVILDSFEPLQAEDEREVSVPAEILRRSRKFPKRAM 1260  
QY 1261 PIWARPDPNPPILLESWKDPDYVPVVGCPPLPAKAPPPIPPRRKRTVVLSESTVSALA 1320  
DB 1261 PIWARPDPNPPILLESWKDPDYVPVVGCPPLPAKAPPPIPPRRKRTVVLSESTVSALA 1320  
QY 1321 ELATKTFGSESAVDSGTATASPDQSDDDAGSDVESYSSMPLEGEGBDPLSDGWS 1380  
DB 1321 ELATKTFGSESAVDSGTATASPDQSDDDAGSDVESYSSMPLEGEGBDPLSDGWS 1380  
QY 1381 STVSEASEDVCCSMYSYTWGALITPCAEEETKLIPINALSNLLRHNILVYATTSRAS 1440  
DB 1381 STVSEASEDVCCSMYSYTWGALITPCAEEETKLIPINALSNLLRHNILVYATTSRAS 1440  
QY 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKLLSVEACKLTTPHARSFYGAK 1500  
DB 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKLLSVEACKLTTPHARSFYGAK 1500  
QY 1501 DVRLNLSKAVNHRSVWKLLEDTETIDITTIHAKNEVFCVQPEKGGKPARLVFPDLG 1560

Db 1501 DVNLSKAVNHIRSVWKKDLLETTETPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 QY 1561 VRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 Db 1561 VRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKQNGCYRRCRASGVLT 1680  
 Db 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKQNGCYRRCRASGVLT 1680  
 QY 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 Db 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 QY 1741 SAPPDPRPKPYDLELITSCSSNVSAHDASGKRVYLYTRDPTPLAANAWEARHTPVN 1800  
 Db 1741 SAPPDPRPKPYDLELITSCSSNVSAHDASGKRVYLYTRDPTPLAANAWEARHTPVN 1800  
 QY 1801 SWLGNIIIMYAPTLWARMILMTHFTSIIILAAQOLEKALDCQIYGACYSLEPLDLQIQL 1860  
 Db 1801 SWLGNIIIMYAPTLWARMILMTHFTSIIILAAQOLEKALDCQIYGACYSLEPLDLQIQL 1860  
 QY 1861 HGLSAFSLHSYSGEINRVASCLRLKGLVPPLRVWHRARSVRARLLSOGGAAATCGKYL 1920  
 Db 1861 HGLSAFSLHSYSGEINRVASCLRLKGLVPPLRVWHRARSVRARLLSOGGAAATCGKYL 1920  
 QY 1921 NWAVRTKLTPTIPAAQOLDLSSWFVAGYSGDDIYHLSLRARPRFWFWCLLLSVGVGIY 1980  
 Db 1921 NWAVRTKLTPTIPAAQOLDLSSWFVAGYSGDDIYHLSLRARPRFWFWCLLLSVGVGIY 1980  
 QY 1981 LLPNR 1985  
 Db 1981 LLPNR 1985

## RESULT 5

ABG30581  
 ID ABG30581 standard; protein; 2201 AA.  
 AC ABG30581;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #1.  
 XX  
 KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;  
 XX cell culture replication; NS2/3; NS3/4; NS3; NS5B.  
 OS Hepatitis C virus.  
 XX  
 PN WO200252015-A2.  
 XX  
 PD 04-JUL-2002.  
 XX  
 PF 20-DEC-2001; 2001WO-CA001843.  
 XX  
 PR 22-DEC-2000; 2000US-0257857P.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.  
 XX  
 PI Kukulj G, Pause A;  
 XX  
 DR WPI; 2002-575382/61.  
 XX  
 DR N-PSDB; ABK88573.  
 XX  
 PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which  
 PT possess enhanced transduction or replication efficiency, useful for  
 PT evaluating potential inhibitors of HCV replication.  
 XX  
 PS Disclosure; Page 49-58; 140pp; English.  
 XX  
 CC The invention describes a self-replicating hepatitis C virus (HCV)

CC polynucleotide molecule comprising a 5'-non translated region (NTR),  
 CC where guanine at position 1 is substituted for adenine, a HCV polyprotein  
 CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-  
 CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating  
 CC potential inhibitors of HCV replication. The HCV RNA molecule is also  
 CC useful for efficiently establishing cell culture replication. The self-  
 CC replicating polynucleotide molecule contains a 5'-NTR, where G at  
 CC position 1 is substituted for A, and therefore provides an alternative to  
 CC existing systems comprising a self-replicating HCV RNA molecule that, in  
 CC conjunction with mutations in the HCV non-structural region, such as the  
 CC G(2042)C/R mutations, transduces and/or replicates with greater  
 CC efficiency. This amino acid sequence is encoded by the hepatitis C virus  
 CC replicon APGK12 and contains the viral protease NS2/3, protease complex  
 CC NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B  
 XX  
 SQ Sequence 2201 AA;

Query Match 100.0%; Score 10462; DB 5; Length 2201;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQQTGRLGCLITSLTGRDRNOVEGEVQVSTATQSFATCNVGVCMVTVHCA 60  
 Db 217 LAPITAYSQQTGRLGCLITSLTGRDRNOVEGEVQVSTATQSFATCNVGVCMVTVHCA 276  
 QY 61 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLVLTTHADVIPVRR 120  
 Db 277 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLVLTTHADVIPVRR 336  
 QY 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
 Db 337 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 396  
 QY 181 RSPVFTDSSPPAVPOTFOVAHLHAPTGSOKSKTKVPAAYAAQGYKVLVLPNSVAATLGF 240  
 Db 397 RSPVFTDSSPPAVPOTFOVAHLHAPTGSOKSKTKVPAAYAAQGYKVLVLPNSVAATLGF 456  
 QY 241 AYMSKAGIDPNIRITGVRIITTCGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
 Db 457 AYMSKAGIDPNIRITGVRIITTCGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 516  
 QY 301 ILIGITVLDQAEATAGARLVVLATATPPGSVTPHPNIEEVALSSTGEIPEYKAIPIETI 360  
 Db 517 ILIGITVLDQAEATAGARLVVLATATPPGSVTPHPNIEEVALSSTGEIPEYKAIPIETI 576  
 QY 361 KGRHILIFCHSKKKCDLAAKLSGLNNAVAYRGLDVSIVPTSGDVIVVATDALTMTGFT 420  
 Db 577 KGRHILIFCHSKKKCDLAAKLSGLNNAVAYRGLDVSIVPTSGDVIVVATDALTMTGFT 636  
 QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSORRGTRGRMGVYRVTPG 480  
 Db 637 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSORRGTRGRMGVYRVTPG 696  
 QY 481 ERPSGMFSSVLCBVDAGCAWYELTPAETSRLRAYLNTPLGVCQDHLFEWESVFTGL 540  
 Db 697 ERPSGMFSSVLCBVDAGCAWYELTPAETSRLRAYLNTPLGVCQDHLFEWESVFTGL 756  
 QY 541 THIDAHFLSQTQAGDNFPYLVAQATVCARAQPPPSWDQWKKLIRLKLPTLHGPTLL 600  
 Db 757 THIDAHFLSQTQAGDNFPYLVAQATVCARAQPPPSWDQWKKLIRLKLPTLHGPTLL 816  
 QY 601 YRLGAVQNEVTTTHPTITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGSVVIVGR 660  
 Db 817 YRLGAVQNEVTTTHPTITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGSVVIVGR 876  
 QY 661 IILSKPAILIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAILGLQTATQKAE 720  
 Db 877 IILSKPAILIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAILGLQTATQKAE 936  
 QY 721 AAPVVESKWRLEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPITTOH 780  
 Db 937 AAPVVESKWRLEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPITTOH 996



QY 61 GSKTLAGKPGITOMYTNVDQDLVQWQAPPGARSILTPCTCGSSDLYLUTRHADVIPVRR 120  
Db 277 GSKTLAGKPGITOMYTNVDQDLVQWQAPPGARSILTPCTCGSSDLYLUTRHADVIPVRR 336  
QY 121 GDSRGLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 337 GDSRGLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 396  
QY 181 RSPVFTDNSSPAPVQCFQVAVHLHAPGSGKSTKVPAAAYAGQYKVLVNLPSVAATLFGF 240  
Db 397 RSPVFTDNSSPAPVQCFQVAVHLHAPGSGKSTKVPAAAYAGQYKVLVNLPSVAATLFGF 456  
QY 241 AYMSKAHGIDNITGVRTITITGAPITVSTYKFLADGCGGGAYDIIICDECHSTDSTT 300  
Db 457 AYMSKAHGIDNITGVRTITITGAPITVSTYKFLADGCGGGAYDIIICDECHSTDSTT 516  
QY 301 ILGTGTVLDOAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEPFGKAIPIETI 360  
Db 517 ILGTGTVLDOAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEPFGKAIPIETI 576  
QY 361 KGGHLLFCHSKKCDLAAKLSGLGLNAVAYRGLDVSVIPTSGDVIWVATDALMTGFT 420  
Db 577 KGGHLLFCHSKKCDLAAKLSGLGLNAVAYRGLDVSVIPTSGDVIWVATDALMTGFT 636  
QY 421 GPDFSVIDCNTVQTQVDFSLDPTFTTETTVPODAVRSORRGTCGRGMGIYRFVTPG 480  
Db 637 GPDFSVIDCNTVQTQVDFSLDPTFTTETTVPODAVRSORRGTCGRGMGIYRFVTPG 696  
QY 481 ERPSGMFDSVLCEDYDAGCAWYELTETPAETSVRLRAYLNTPGLPVCDHLEFVESVFTGL 540  
Db 697 ERPSGMFDSVLCEDYDAGCAWYELTETPAETSVRLRAYLNTPGLPVCDHLEFVESVFTGL 756  
QY 541 THIDAFLSQTQAGDNFPLVAYQATVCARAQAPPSWDMKCLRLKPTLHGPTPLL 600  
Db 757 THIDAFLSQTQAGDNFPLVAYQATVCARAQAPPSWDMKCLRLKPTLHGPTPLL 816  
QY 601 YRLGAVQNEVTHPTITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSGVIVGR 660  
Db 817 YRLGAVQNEVTHPTITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSGVIVGR 876  
QY 661 IILSGKPAIIPREVLRYBDEMEECASHLPYIEQGMQLAEQFKQKAIQLLOTATQKQAE 720  
Db 877 IILSGKPAIIPREVLRYBDEMEECASHLPYIEQGMQLAEQFKQKAIQLLOTATQKQAE 936  
QY 721 AAPVVESKWRTEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITPLTQH 780  
Db 937 AAPVVESKWRTEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITPLTQH 996  
QY 781 TLLFNILGWVAAQLAPPASAASAFVAGIAGAAVGSIGLGVLDLILAGVGAGVAGALVA 840  
Db 997 TLLFNILGWVAAQLAPPASAASAFVAGIAGAAVGSIGLGVLDLILAGVGAGVAGALVA 1056  
QY 841 FKVMGEMSTEDLNLNLLPAILSPGALVGVVCAAILRRHVGFEGEAVQWNRLLIAFASR 900  
Db 1057 FKVMGEMSTEDLNLNLLPAILSPGALVGVVCAAILRRHVGFEGEAVQWNRLLIAFASR 1116  
QY 901 GNHVSPTHVPSDAAARVTOILSSITITQLKRLHWNEDCSTPCSGSMLRDVMDWIC 960  
Db 1117 GNHVSPTHVPSDAAARVTOILSSITITQLKRLHWNEDCSTPCSGSMLRDVMDWIC 1176  
QY 961 TVLTDFKTLQSKLLPRLPGVPFFSCQGYKGVWRGDGMOTTCPCGAQITGHVKNXGMR 1020  
Db 1177 TVLTDFKTLQSKLLPRLPGVPFFSCQGYKGVWRGDGMOTTCPCGAQITGHVKNXGMR 1236  
QY 1021 IVGPRTCSNTHGTPIINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDFFHYVTQM 1080  
Db 1237 IVGPRTCSNTHGTPIINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDFFHYVTQM 1296  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRHLRHYPACKPLIREVTVLGNLQYLVGSQPCPE 1140  
Db 1297 TTDNVKCPQVPAPEFFTEVDGVRHLRHYPACKPLIREVTVLGNLQYLVGSQPCPE 1356

QY 1141 PDVAVLTSMLTDPGHITAEAKRIARGSPPSLASASSASQLSAKTCTTRHDSPAD 1200  
Db 1357 PDVAVLTSMLTDPGHITAEAKRIARGSPPSLASASSASQLSAKTCTTRHDSPAD 1416  
QY 1201 LIEANLWRQMGNGNITRVESENKVWILDSPEPLQAEDEDEREVSPAEILRRSRKFFPAM 1260  
Db 1417 LIEANLWRQMGNGNITRVESENKVWILDSPEPLQAEDEDEREVSPAEILRRSRKFFPAM 1476  
QY 1261 PIWARPNDYNPLLESWKDPDYPVPVHGCPLPPAKAPEIPPPRRKRTVVLESTVSSALA 1320  
Db 1477 PIWARPNDYNPLLESWKDPDYPVPVHGCPLPPAKAPEIPPPRRKRTVVLESTVSSALA 1536  
QY 1321 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPLEGECPDPLSDGSW 1380  
Db 1537 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPLEGECPDPLSDGSW 1596  
QY 1381 STVSEEEASEDVCSSMSYTWGALITPCAABETKLPINALSNSLLRRHNLVYATTSRAS 1440  
Db 1597 STVSEEEASEDVCSSMSYTWGALITPCAABETKLPINALSNSLLRRHNLVYATTSRAS 1656  
QY 1441 LRQKKVTFDLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEFGYAK 1500  
Db 1657 LRQKKVTFDLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEFGYAK 1716  
QY 1501 DVNLSKSAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
Db 1717 DVNLSKSAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 1776  
QY 1561 VRVCEKMALDYVVTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 1620  
Db 1777 VRVCEKMALDYVVTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 1836  
QY 1621 TVTENDIRVEESYIQCCLDAPPEARQAIIRSLTERLYTGGPLTNSKGONCYRRCPASVLT 1880  
Db 1837 TVTENDIRVEESYIQCCLDAPPEARQAIIRSLTERLYTGGPLTNSKGONCYRRCPASVLT 1896  
QY 1681 TSCGNTLTCLYKAAAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEDEASLRAFTTEAMTRY 1740  
Db 1897 TSCGNTLTCLYKAAAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEDEASLRAFTTEAMTRY 1956  
QY 1741 SAPPGGPPKPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPLARAWEARHTFPVN 1800  
Db 1957 SAPPGGPPKPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPLARAWEARHTFPVN 2016  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
Db 2017 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIQRL 2076  
QY 1861 HGLSAPSLHSYSPGEINRVASCLRLKLGVPPLRVWRHRSVRARLLSQGGAATCGKLYF 1920  
Db 2077 HGLSAPSLHSYSPGEINRVASCLRLKLGVPPLRVWRHRSVRARLLSQGGAATCGKLYF 2136  
QY 1921 NNAVRTKLLTPIPAASQDLSSWVAGYSGGDIYHLSLRARPRPMFWCMLLLSVGVGIY 1980  
Db 2137 NNAVRTKLLTPIPAASQDLSSWVAGYSGGDIYHLSLRARPRPMFWCMLLLSVGVGIY 2196  
QY 1981 LLPNR 1985  
Db 2197 LLPNR 2201

## RESULT 7

ABG32451

ID ABG32451 standard; protein; 3010 AA.

XX AC ABG32451;

XX AC ABG32451;

DT 15-NOV-2002 (first entry)

DE Hepatitis C virus Con 1 isolate polyprotein.

XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;



internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.  
Hepatitis C virus.  
WO200259321-A2.  
01-AUG-2002.  
16-JAN-2002; 2002WO-EP000526.  
23-JAN-2001; 2001US-0263479P.  
(RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
De Francesco R, Migliaccio G, Paonessa G;  
WPI; 2002-599793/64.  
N-PSDB; ABK91411.  
New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
ribosome entry site (IRES) region, useful in studying HCV replication and  
expression.  
Claim 1; Page 34-36; 69pp; English.  
The invention relates to nucleic acid molecules comprising altered HCV  
NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
internal ribosome entry site (IRES) region coding for one or more NS3,  
NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
are detailed in the specification. Also included are (1) an expression  
vector comprising a nucleotide sequence coding for the altered nucleic  
acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
recombinant cell human hepatoma cell comprising the altered nucleic acids  
; (3) a recombinant cell produced by introducing into a human hepatoma  
cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
replicon enhanced cell or which containing a functional HCV replicon; (5)  
an HCV replicon enhanced cells made in the method; and (6) measuring the  
ability of a compound to affect HCV activity. The HCV replicons and HCV  
replicon enhanced cells are useful in studying HCV replication and  
expression, and HCV and host cell interactions, producing HCV RNA and  
proteins, and providing a system for measuring the ability of a compound  
to modulate one or more HCV activities e.g. to discover drugs which may  
treat HCV mediated diseases such as liver failure, cirrhosis and  
hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
and NS5B proteins) used as a basis for the adaptive mutations of the  
invention

Sequence 3010 AA;  
Query Match 100.0%; Score 10462; DB 5; Length 3010;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
1 MAPITAYSOQTRGLLCITSLTGRDRNQVEGEVQVSTATQSFATCNGVCWTVYHGA 60  
1026 LAPITAYSOQTRGLLCITSLTGRDRNQVEGEVQVSTATQSFATCNGVCWTVYHGA 1085  
61 GSKTLGAPKGPITOMYTNVDQDLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 120  
1086 GSKTLGAPKGPITOMYTNVDQDLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 1145  
121 GDSRGSLLSPRPVSYLKGSGGPLLCPSSHAGVIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
1146 GDSRGSLLSPRPVSYLKGSGGPLLCPSSHAGVIFRAAVCTRGVAKAVDFVPVSEMETTM 1205  
181 RSPVFTDSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAAQYKVLVLPNSVAATLGF 240  
1206 RSPVFTDSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAAQYKVLVLPNSVAATLGF 1265  
241 AYMSKAHGIDPNIRKTGVRTITTGAPITYSTYVGKFLADGGCGGGAYDIIICDECHSTDTT 300

Db 1266 AYMSKAHGIDPNIRKTGVRTITTGAPITYSTYVGKFLADGGCGGGAYDIIICDECHSTDTT 1325  
QY 301 ILGIGTVLDOAETAGARLVVLAATATPPGSGTVVPHNIEEVALSTGEIPFYGKAIPBETI 360  
Db 1326 ILGIGTVLDOAETAGARLVVLAATATPPGSGTVVPHNIEEVALSTGEIPFYGKAIPBETI 1385  
QY 361 KGRHLIFCHSKKCKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIIVVATDALMTGFT 420  
Db 1386 KGRHLIFCHSKKCKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIIVVATDALMTGFT 1445  
QY 421 GDFSVIDCNCVCTQTVDFSLDPTFTTETTVPODAVRSORRGRTGRGMGIYRFVTPPG 480  
Db 1446 GDFSVIDCNCVCTQTVDFSLDPTFTTETTVPODAVRSORRGRTGRGMGIYRFVTPPG 1505  
QY 481 ERPSGMFDSVLCYDAGCAWYELTTPAETSRLRAYLNTPLPVCQDHLFWESVFTGL 540  
Db 1506 ERPSGMFDSVLCYDAGCAWYELTTPAETSRLRAYLNTPLPVCQDHLFWESVFTGL 1565  
QY 541 THIDAHFLSOTKQAGDNPPYLVAQATVCARAQAPPPSDQWKKLILKPLTHGPTPLL 600  
Db 1566 THIDAHFLSOTKQAGDNPPYLVAQATVCARAQAPPPSDQWKKLILKPLTHGPTPLL 1625  
QY 601 YRLGAVQNEVTTTHPTIKYINACHMSADLEVVTSTWLVGGVLAALAAAYCLITGSGVIVGR 660  
Db 1626 YRLGAVQNEVTTTHPTIKYINACHMSADLEVVTSTWLVGGVLAALAAAYCLITGSGVIVGR 1685  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQALAEQKAKIIGLOTATQABA 720  
Db 1686 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQALAEQKAKIIGLOTATQABA 1745  
QY 721 AAPVYESKWRITLAEAWKHMNFISGQYLAGSLTLPNPAIASMAFTASITSLTTOH 780  
Db 1746 AAPVYESKWRITLAEAWKHMNFISGQYLAGSLTLPNPAIASMAFTASITSLTTOH 1805  
QY 781 TLLFNILGWAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 840  
Db 1806 TLLFNILGWAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 1865  
QY 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVMMRLIAFASR 900  
Db 1866 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVMMRLIAFASR 1925  
QY 901 GNVHPTVPSDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSSWLRDWDWIC 960  
Db 1926 GNVHPTVPSDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSSWLRDWDWIC 1985  
QY 961 TVLTDFKTLQSKLLPRLPGVPPFFSCQYKGVWRGDMQITTCPCGAQITGHVKNGSMR 1020  
Db 1986 TVLTDFKTLQSKLLPRLPGVPPFFSCQYKGVWRGDMQITTCPCGAQITGHVKNGSMR 2045  
QY 1021 IVGPRTCSTNTHGTFPPINAYTTGCTPPSPAPNYSRALWRVAABEYVEVTRVGDHFHYVTGM 1080  
Db 2046 IVGPRTCSTNTHGTFPPINAYTTGCTPPSPAPNYSRALWRVAABEYVEVTRVGDHFHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLREEVTVLGVNLQVLVGSQLPCEPE 1140  
Db 2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLREEVTVLGVNLQVLVGSQLPCEPE 2165  
QY 1141 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTTHDSDPAD 1200  
Db 2166 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTTHDSDPAD 2225  
QY 1201 LIEANLLWRQEMGNITRVESENKVLDSFEPLOABEDEREVSVAEILRRSRKEPRAM 1260  
Db 2226 LIEANLLWRQEMGNITRVESENKVLDSFEPLOABEDEREVSVAEILRRSRKEPRAM 2285  
QY 1261 PIWARDYNPPLLESWKDPDYVPVVGCPPLPAKAPPIPPPRKRKRWLSESTVSSALA 1320  
Db 2286 PIWARDYNPPLLESWKDPDYVPVVGCPPLPAKAPPIPPPRKRKRWLSESTVSSALA 2345  
QY 1321 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEFGDPLDLSGWS 1380  
Db 2346 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEFGDPLDLSGWS 2405



```

QY 1381 STVSEASEDVVCCSMSTWTGALITPCAAETKLPINALSNLSLRHNLVYATTSRSAS 1440
Db 2406 STVSEASEDVVCCSMSTWTGALITPCAAETKLPINALSNLSLRHNLVYATTSRSAS 2465
QY 1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGA 1500
Db 2466 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGA 2525
QY 1501 DVRLNSSKAVNHRSVWKDLLEDTEPTDTTMAKNEVFCVQPEKGRKPARLIIVFPDLG 1560
Db 2526 DVRLNSSKAVNHRSVWKDLLEDTEPTDTTMAKNEVFCVQPEKGRKPARLIIVFPDLG 2585
QY 1561 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 1620
Db 2586 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 2645
QY 1621 TVTENDIRVESIYQCCDLAPEARQATRSRLTERLYIGGFLTNSKGQNGYRRCRASGVLT 1680
Db 2646 TVTENDIRVESIYQCCDLAPEARQATRSRLTERLYIGGFLTNSKGQNGYRRCRASGVLT 2705
QY 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTEAMTRY 1740
Db 2706 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTEAMTRY 2765
QY 1741 SAPGDPKPEYDLELITSCSSNVSVADHASKGRVYVLTTRDPTTPLARAAWETAHTPVN 1800
Db 2766 SAPGDPKPEYDLELITSCSSNVSVADHASKGRVYVLTTRDPTTPLARAAWETAHTPVN 2825
QY 1801 SWLGNIIMYAPTLLWARMILMTHFFSILLAQEKLDCQIYGACYSIEPLDLPQIQLR 1860
Db 2826 SWLGNIIMYAPTLLWARMILMTHFFSILLAQEKLDCQIYGACYSIEPLDLPQIQLR 2885
QY 1861 HGLSAPSLHSYSGEINRVASCLRKLGVPPLVWRHRSVRARLLSQGGRATCGKYL 1920
Db 2886 HGLSAPSLHSYSGEINRVASCLRKLGVPPLVWRHRSVRARLLSQGGRATCGKYL 2945
QY 1921 NWAVRTKLTLPAPASQLDSSWFVAGYSGGDIYHLSLRARPRFWMCULLLSVGVGIY 1980
Db 2946 NWAVRTKLTLPAPASQLDSSWFVAGYSGGDIYHLSLRARPRFWMCULLLSVGVGIY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 8
AAE15731
ID AAE15731 standard; protein; 1985 AA.
XX AC AAE15731;
XX DT 12-MAR-2002 (first entry)
XX DE Hepatitis C virus (HCV) rep1bBartMan polyprotein variant S1172P.
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW rep1bBartMan; mutant; mutein; variant.
OS Hepatitis C virus.
OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 1172 /note= "wild type Ser substituted with Pro"
XX PN WO200189364-A2.
XX PD 29-NOV-2001.
XX PF 23-MAY-2001; 2001WO-US016822.
XX

```

```

PR 23-MAY-2000; 2000US-00576989.
XX (UNITW ) UNIV WASHINGTON.
XX PA Rice CM, Blight KJ;
XX PI WPI; 2002-066755/09.
XX DR Hepatitis C virus variants having greater transfection efficiency and
XX PT ability to survive subpassage, useful as a vaccine for immunizing primate
XX PT to the virus, comprise non-naturally occurring viral sequences.
XX PS Claim 14; Page; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated HCV
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) rep1bBartMan polyprotein
CC variant. Note: The present sequence is not shown in the specification but
CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
CC specification
XX
SQ Sequence 1985 AA;
Query Match 100.0%; Score 10460; DB 5; Length 1985;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAPITAYSOOTRGLGCIITSLTGRDNOVEGEVQVWSTATOSFLATCNGVCWTYHGA 60
Db 1 MAPITAYSOOTRGLGCIITSLTGRDNOVEGEVQVWSTATOSFLATCNGVCWTYHGA 60
QY 61 GSKTLAGPKGPITOMYTNVDQDLVQWQAPPAGARSILTPCTCGSSDLYLVTRHADVIPRRR 120
Db 61 GSKTLAGPKGPITOMYTNVDQDLVQWQAPPAGARSILTPCTCGSSDLYLVTRHADVIPRRR 120
QY 121 GDSRGSLLSPRPVSYLKGSSGGLLCPSGHVGIFPRAAVCTRGVAKAVDFVPVESMETM 180
Db 121 GDSRGSLLSPRPVSYLKGSSGGLLCPSGHVGIFPRAAVCTRGVAKAVDFVPVESMETM 180
QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTGSKSTKVPAAAGYKVLVNLNPSVAATLFG 240
Db 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTGSKSTKVPAAAGYKVLVNLNPSVAATLFG 240
QY 241 AYMSKAHGIDNIRTVRTITTGAPITYSTYKFLADGCGSGGAYDIICDECHSTDSTT 300
Db 241 AYMSKAHGIDNIRTVRTITTGAPITYSTYKFLADGCGSGGAYDIICDECHSTDSTT 300
QY 301 ILGITVLDQAEATAGARLWLATATPPGTVTPHNEIEVALSSTGEIPFYGKAIPETI 360

```

301	ILGIGTVLDOAETAGARLVUATATPPGSSVTPHPNIEEVALSGTGHIPFYGVKAIPIETI	360	1441	LROKKVTFDRILQVLDHHRDVLKEMKAKASTVKAKLLSVEACKLTPPHSARSKEFGYCAK	1500
Db			QY		
361	KGGRHLIFCHSKKCCDELAALSLGLGNANAYRGLDVSVIPTSGDVIIVATDALMTGFT	420	1441	LROKKVTFDRILQVLDHHRDVLKEMKAKASTVKAKLLSVEACKLTPPHSARSKEFGYCAK	1500
QY			Db		
361	KGGRHLIFCHSKKCCDELAALSLGLGNANAYRGLDVSVIPTSGDVIIVATDALMTGFT	420	1501	DVRNLSSKANVHRSWKDLLEDTEPIDITIMAKNEVFCVQPEKGGKPKPARLLIVFPDLG	1560
Db			QY		
421	GDPSVIDCNTCVTQVDFSLDPTFTTETTVPODAVSRSORRGTRGMRGIVREFTPG	480	1501	DVRNLSSKANVHRSWKDLLEDTEPIDITIMAKNEVFCVQPEKGGKPKPARLLIVFPDLG	1560
QY			Db		
421	GDPSVIDCNTCVTQVDFSLDPTFTTETTVPODAVSRSORRGTRGMRGIVREFTPG	480	1561	VRCEKVALYDVSTLPOAVMGSSYGFQYSPGQVRFVLVNAWKAKKCPMGFAYDTRCFDS	1620
Db			QY		
481	BRPSGMPDSSVLCCEYDAGCAWYELTPAETSVRILRAYLNTPLPVCQDHLFEWESVFTGL	540	1561	VRCEKVALYDVSTLPOAVMGSSYGFQYSPGQVRFVLVNAWKAKKCPMGFAYDTRCFDS	1620
QY			Db		
481	BRPSGMPDSSVLCCEYDAGCAWYELTPAETSVRILRAYLNTPLPVCQDHLFEWESVFTGL	540	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGGLTNSKQNGCYRRCRASGVLT	1680
Db			QY		
541	THIDAHFELSQTQAGDNFPYLVAQYATVCARAQAPPSPWDQWKCLIRLKTPLHGTPL	600	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGGLTNSKQNGCYRRCRASGVLT	1680
QY			Db		
541	THIDAHFELSQTQAGDNFPYLVAQYATVCARAQAPPSPWDQWKCLIRLKTPLHGTPL	600	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTAMTRY	1740
Db			QY		
601	YRLGAVQNEVTTTHPITIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	660	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTAMTRY	1740
QY			Db		
601	YRLGAVQNEVTTTHPITIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	660	1741	SAPPGDPKPEVDLELITSCSNVSVAHDAHGKRVYVLTTRDPTTPLARAAMETARHTPN	1800
Db			QY		
661	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQKALGLLQATKQAEA	720	1741	SAPPGDPKPEVDLELITSCSNVSVAHDAHGKRVYVLTTRDPTTPLARAAMETARHTPN	1800
QY			Db		
661	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQKALGLLQATKQAEA	720	1801	SWLGNIMVAPTLWARMILMTHFFSILLAQEQLEKALDQIYGACYSIEPLDLPQIQL	1860
Db			QY		
721	AAPVVEKWRITLFAFWKHMNFISGIVLAGLSTLPGNPATASLMAFTASITSLTTOH	780	1801	SWLGNIMVAPTLWARMILMTHFFSILLAQEQLEKALDQIYGACYSIEPLDLPQIQL	1860
QY			Db		
721	AAPVVEKWRITLFAFWKHMNFISGIVLAGLSTLPGNPATASLMAFTASITSLTTOH	780	1861	HGLSAFSLHSYSPGINRVASCLRGVPPPLRVWRHARSVRARLLSQGGRATCGKYL	1920
Db			QY		
781	TLLFNILGWAAQLAPPSSAFAFGAGIAGAGVSGIGLKVLDLILAGYGAGVALVA	840	1861	HGLSAFSLHSYSPGINRVASCLRGVPPPLRVWRHARSVRARLLSQGGRATCGKYL	1920
QY			Db		
781	TLLFNILGWAAQLAPPSSAFAFGAGIAGAGVSGIGLKVLDLILAGYGAGVALVA	840	1921	NWAVRTKLKLTPIPAASQDLSSWVAGVSGGDIYHLSRARPRFMWCLLLSVGVGIY	1980
Db			QY		
841	EKVMSEMPSTEDLNNLLPAIILSPGALVGVVCAAILRRHVGGEGAVQWMLRIAFASR	900	1921	NWAVRTKLKLTPIPAASQDLSSWVAGVSGGDIYHLSRARPRFMWCLLLSVGVGIY	1980
QY			Db		
841	EKVMSEMPSTEDLNNLLPAIILSPGALVGVVCAAILRRHVGGEGAVQWMLRIAFASR	900	1981	LLPNR 1985	
Db			QY		
901	GNHVSPTHVPESDAAARVQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960	1981	LLPNR 1985	
QY			Db		
901	GNHVSPTHVPESDAAARVQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960	RESULT 9		
QY			AAE15730		
961	TVLTDFTKWLQSKLLPRLPVPFSCORGKVGWGDGMQTTCPGQAQITGHVKNGSMR	1020	ID	AAE15730 standard; protein; 1985 AA.	
QY			XX		
961	TVLTDFTKWLQSKLLPRLPVPFSCORGKVGWGDGMQTTCPGQAQITGHVKNGSMR	1020	AC	AAE15730;	
Db			XX		
1021	IVGPRICSNTHGTFPPINATYTGCTPSPAPNYSRALRVAAEYVEVTRVGDHYVTGM	1080	XX	12-MAR-2002 (first entry)	
QY			XX		
1021	IVGPRICSNTHGTFPPINATYTGCTPSPAPNYSRALRVAAEYVEVTRVGDHYVTGM	1080	DE	Hepatitis C virus (HCV) repibBartMan polyprotein variant S1172C.	
Db			XX		
1081	TTDNVCKPCQVPAPEFTEVDGVRHLRYAPACKPILLREEVTFVLGNQYLVGSQLPCEPE	1140	XX	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;	
QY			XX		
1081	TTDNVCKPCQVPAPEFTEVDGVRHLRYAPACKPILLREEVTFVLGNQYLVGSQLPCEPE	1140	KW	gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;	
Db			KW	repibBartMan; mutant; mutein; variant.	
1141	PDVAULTSMLTDPHSHTAETAKRELARGSPPLASSASQUSAPSLKATCTTRHDSPDAD	1200	XX	Hepatitis C virus.	
QY			OS	Synthetic.	
1141	PDVAULTSMLTDPHSHTAETAKRELARGSPPLASSASQUSAPSLKATCTTRHDSPDAD	1200	XX		
Db			XX		
1201	LIEANLLWRQBMGNGNITRVESENKVIILDSPEPLQAEDEDEREVSVPAILRRSRKFFRAM	1260	XX	Key	
QY			XX	Location/Qualifiers	
1201	LIEANLLWRQBMGNGNITRVESENKVIILDSPEPLQAEDEDEREVSVPAILRRSRKFFRAM	1260	XX	Misc-difference 1172	
Db			XX	/note= "Wild type Ser substituted with Cys"	
1261	PIWARPDPNPLLESWKDPDVPVPHGCPPLPPAKAPPIPPPRKRTVILSESTVSSALA	1320	XX	WO200189364-A2.	
QY			XX	29-NOV-2001.	
1261	PIWARPDPNPLLESWKDPDVPVPHGCPPLPPAKAPPIPPPRKRTVILSESTVSSALA	1320	XX	23-MAY-2001; 2001WO-US016822.	
Db			XX	23-MAY-2000; 2000US-00576989.	
1321	ELATKTFGSSSSAVDSSTATASPDQDDGAGSDVESYSPMPLEGECPDPLSDGSW	1380	XX	(UNIW ) UNIV WASHINGTON.	
QY			XX		
1321	ELATKTFGSSSSAVDSSTATASPDQDDGAGSDVESYSPMPLEGECPDPLSDGSW	1380	XX		
Db			XX		
1381	STVSEASESDVVCSSMSYTWGALITPCAABETKLPINALNSLLRHHNLVYATTSRAS	1440	XX		
QY			XX		
1381	STVSEASESDVVCSSMSYTWGALITPCAABETKLPINALNSLLRHHNLVYATTSRAS	1440	XX		
Db			XX		

PI Rice CM, Blight KJ;  
 DR WPI; 2002-066755/09.  
 XX  
 XX Hepatitis C virus variants having greater transfection efficiency and  
 PT ability to survive subpassage, useful as a vaccine for immunizing primate  
 PT to the virus, comprise non-naturally occurring viral sequences.  
 XX  
 XX Claim 14; Page; 174pp; English.  
 XX  
 CC The invention relates to Hepatitis C virus (HCV) variants which include  
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
 CC variants that have a transfection efficiency and ability to survive  
 CC subpassage greater than HCV that have wild-type polypeptide coding  
 CC regions. The polynucleotides of the invention are useful for identifying  
 CC a cell line that is permissive for infection with HCV and detecting  
 CC replication of HCV in cells of the cell line. They are also useful for  
 CC testing a compound for anti-viral properties and for inhibiting HCV  
 CC infection. They are also useful for the generation of defined HCV virus  
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
 CC attachment, penetration and entry, structure/function studies on HCV  
 CC proteins and RNA elements and identification of new antiviral targets, a  
 CC systematic survey of cell culture systems and conditions to identify  
 CC those that support wild-type and variant HCV RNA replication and particle  
 CC release, production of adaptive HCV variants capable of more efficiency  
 CC replication in cell culture, production of HCV variants with altered  
 CC tissue or species tropism, establishment of alternative animal models for  
 CC inhibitor evaluation including those supporting HCV variant replication,  
 CC development of cell-free HCV replication assays, production of  
 CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
 CC derivatives as possible vaccine candidates, engineering of attenuated or  
 CC defective HCV derivatives for expression of heterologous gene products  
 CC for gene therapy and vaccine applications and for utilisation of the HCV  
 CC glycoproteins for targeted delivery of therapeutic agents to the liver  
 CC or other cell types with appropriate receptors. Vaccine comprising these  
 CC sequences is useful for inducing immunoprotection to HCV in a primate.  
 CC The present sequence is Hepatitis C virus (HCV) repBartMan polypeptide  
 CC variant. Note: The present sequence is not shown in the specification but  
 CC is derived from SEQ ID NO: 3 (AAE15171) shown in page 65 of the  
 CC specification  
 XX  
 SQ Sequence 1985 AA;  
 Query Match 100.0%; Score 10460; DB 5; Length 1985;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGVEQVQVSTATQSLATCVNGVCWTVYHGA 60  
 DB 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGVEQVQVSTATQSLATCVNGVCWTVYHGA 60  
 QY 61 GSKTLGAPKGPITOMYTNVDQDLVGVQAPPGARSLTPTCGSSDLVLRHADVIPVRRR 120  
 DB 61 GSKTLGAPKGPITOMYTNVDQDLVGVQAPPGARSLTPTCGSSDLVLRHADVIPVRRR 120  
 QY 121 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRCVAKADVFPVPSMETTM 180  
 DB 121 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRCVAKADVFPVPSMETTM 180  
 QY 181 RSPVFTDSSPPAVPOTFOVHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240  
 DB 181 RSPVFTDSSPPAVPOTFOVHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240  
 QY 241 AYMSKAHGDINRTGVRITTCAPITTYTYGKFLADGCGSGGAYDIIICDECHSTDSTT 300  
 DB 241 AYMSKAHGDINRTGVRITTCAPITTYTYGKFLADGCGSGGAYDIIICDECHSTDSTT 300  
 QY 301 ILGIGTVLQDQATAGARLVVLTATPPGVTVPHPNI EEVALSSTGEIIFYGKAIPETI 360  
 DB 301 ILGIGTVLQDQATAGARLVVLTATPPGVTVPHPNI EEVALSSTGEIIFYGKAIPETI 360  
 QY 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIWVATDALMTGFT 420

Db 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIWVATDALMTGFT 420  
 QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAVSRORSORRGTRGRMGIVFVFTPG 480  
 Db 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAVSRORSORRGTRGRMGIVFVFTPG 480  
 QY 481 ERPSGMDSSVLCBCEYDAGCAWYELTTPAETSVRLRAYLNTFGLPVCQDHLFEWESVFTGL 540  
 Db 481 ERPSGMDSSVLCBCEYDAGCAWYELTTPAETSVRLRAYLNTFGLPVCQDHLFEWESVFTGL 540  
 QY 541 THIDAHFLSOTKOAGDNFPYLVAYOATVCARAQAPPSWDQWKLRLKLTPLHGTPTLL 600  
 Db 541 THIDAHFLSOTKOAGDNFPYLVAYOATVCARAQAPPSWDQWKLRLKLTPLHGTPTLL 600  
 QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGVIWGR 660  
 Db 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGVIWGR 660  
 QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKOKAIGLLQTATKQAE 720  
 Db 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKOKAIGLLQTATKQAE 720  
 QY 721 AAPVSEKWTLEAFWAKHWNFIISGLIYLAGLSTLPGNPAIASLMAFTASITSLTQH 780  
 Db 721 AAPVSEKWTLEAFWAKHWNFIISGLIYLAGLSTLPGNPAIASLMAFTASITSLTQH 780  
 QY 781 TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
 Db 781 TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
 QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 Db 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 QY 901 GNVSPHYVPESDAARVTQILSSLTITQLKRLHWNEDCSTPCSGSWLRDWDWIC 960  
 Db 901 GNVSPHYVPESDAARVTQILSSLTITQLKRLHWNEDCSTPCSGSWLRDWDWIC 960  
 QY 961 TVLTDFKTLQSKLPLRLPGVPFFSCQYGVWVRGDMQTTCPGCAQITGHVKGSMR 1020  
 Db 961 TVLTDFKTLQSKLPLRLPGVPFFSCQYGVWVRGDMQTTCPGCAQITGHVKGSMR 1020  
 QY 1021 IVGPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALRVAAEYVEVTRVGDHYVTGM 1080  
 Db 1021 IVGPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALRVAAEYVEVTRVGDHYVTGM 1080  
 QY 1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKELLREEVTLVGLNQYLVGSQLPCEPE 1140  
 Db 1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKELLREEVTLVGLNQYLVGSQLPCEPE 1140  
 QY 1141 PDVAVLTSLMTDPSHITAETAKRLARGSPPSLASSASSQLSAPSLKATCTTRHDSPAD 1200  
 Db 1141 PDVAVLTSLMTDPSHITAETAKRLARGSPPSLASSASSQLSAPSLKATCTTRHDSPAD 1200  
 QY 1201 LIEANLLMRQEGNITRVESENKVVILDSFPELQAEDEREVSPAEILRRSRKFPFRAM 1260  
 Db 1201 LIEANLLMRQEGNITRVESENKVVILDSFPELQAEDEREVSPAEILRRSRKFPFRAM 1260  
 QY 1261 PIWARPDPNPLLESWKOPDYVPPVHGCPLPAKAPPIPPRRKRKTVVLSSTVSSALA 1320  
 Db 1261 PIWARPDPNPLLESWKOPDYVPPVHGCPLPAKAPPIPPRRKRKTVVLSSTVSSALA 1320  
 QY 1321 ELATKTFGSSSESSAVDSGTATASPPQSDGSDGSDVESYSSMPPLEGEPPGDDSDGSM 1380  
 Db 1321 ELATKTFGSSSESSAVDSGTATASPPQSDGSDGSDVESYSSMPPLEGEPPGDDSDGSM 1380  
 QY 1381 STVSEASEDVVCCSMTWTGALTIPCAAEETKLPINALNSLLRHHNLVYATTSRSAS 1440  
 Db 1381 STVSEASEDVVCCSMTWTGALTIPCAAEETKLPINALNSLLRHHNLVYATTSRSAS 1440  
 QY 1441 LRQKVTVDRLQVLDHDDYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 1500  
 Db 1441 LRQKVTVDRLQVLDHDDYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 1500

QY	1501	DVRNLSKAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560	PT	
Db	1501	DVRNLSKAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560	PT	
QY	1561	VRCEKALYDVVSTLPQAVMGSSYGFQYSGQRFVEFLVNAWKAKKCPMGAYDTRCFDS	1620	XX	
Db	1561	VRCEKALYDVVSTLPQAVMGSSYGFQYSGQRFVEFLVNAWKAKKCPMGAYDTRCFDS	1620	XX	
QY	1621	TVTENDIRVESIYOCCLAPAPARQAIRSLTERLVIGGPLTNSKQNGCYRRCRASGVLT	1680	PS	
Db	1621	TVTENDIRVESIYOCCLAPAPARQAIRSLTERLVIGGPLTNSKQNGCYRRCRASGVLT	1680	PS	
QY	1681	TSCGNLTTCYLKAAACRAALQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY	1740	CC	
Db	1681	TSCGNLTTCYLKAAACRAALQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY	1740	CC	
QY	1741	SAPPGDPKPEYDLELITSCSSNVSAHDASGRVYVYLTRDPTTTLARAAMETARHTPVN	1800	CC	
Db	1741	SAPPGDPKPEYDLELITSCSSNVSAHDASGRVYVYLTRDPTTTLARAAMETARHTPVN	1800	CC	
QY	1801	SWLGNIMVAPTLWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL	1860	CC	
Db	1801	SWLGNIMVAPTLWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL	1860	CC	
QY	1861	HGLSAFSLHSYSPGEINRVASCLRKLGVPPPLRVWHRARSVRARLLSQGGRAATCGKLYF	1920	CC	
Db	1861	HGLSAFSLHSYSPGEINRVASCLRKLGVPPPLRVWHRARSVRARLLSQGGRAATCGKLYF	1920	CC	
QY	1921	NNAVTKLKLTPAASQDLSSWFVAGYSGGDIYHLSLRPRPFMWCLLLLSVGVGIY	1980	CC	
Db	1921	NNAVTKLKLTPAASQDLSSWFVAGYSGGDIYHLSLRPRPFMWCLLLLSVGVGIY	1980	CC	
QY	1981	LLPNR 1985		CC	
Db	1981	LLPNR 1985		CC	
RESULT 10					XX
AAE15727					XX
ID	AAE15727 standard; protein; 1985 AA.				XX
AC	AAE15727;				XX
DT	12-MAR-2002 (first entry)				XX
DE	Hepatitis C virus (HCV) replBartMan polyprotein variant S11791.				XX
KW	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;				XX
KW	gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;				XX
KW	replBartMan; mutant; mutein; variant.				XX
OS	Hepatitis C virus.				XX
OS	Synthetic.				XX
Key	Location/Qualifiers				XX
FT	Misc-difference 1179				XX
FT	/note= "Wild type Ser substituted with Ile"				XX
XX	W0200189364-A2.				XX
XX	29-NOV-2001.				XX
XX	23-MAY-2001; 2001W0-US016822.				XX
XX	23-MAY-2000; 2000US-00576989.				XX
XX	(UNIM ) UNIV WASHINGTON.				XX
XX	Rice CM, Blight KJ;				XX
XX	WPI; 2002-066755/09.				XX

Hepatitis C virus variants having greater transfection efficiency and ability to survive subpassage, useful as a vaccine for immunizing primate to the virus, comprise non-naturally occurring viral sequences.

Claim 14; Page; 174pp; English.

The invention relates to Hepatitis C virus (HCV) variants which include polynucleotides comprising non-naturally occurring HCV sequence and HCV variants that have a transfection efficiency and ability to survive subpassage greater than HCV that have wild-type polyprotein coding regions. The polynucleotides of the invention are useful for identifying a cell line that is permissive for infection with HCV and detecting replication of HCV in cells of the cell line. They are also useful for testing a compound for anti-viral properties and for inhibiting HCV infection. They are also useful for the generation of defined HCV virus stocks to develop in vitro and in vivo assays for virus neutralisation, attachment, penetration and entry, structure/function studies on HCV proteins and RNA elements and identification of new antiviral targets, a systematic survey of cell culture systems and conditions to identify those that support wild-type and variant HCV RNA replication and particle release, production of adaptive HCV variants capable of more efficient replication in cell culture, production of HCV variants with altered tissue or species tropism, establishment of alternative animal models for inhibitor evaluation including those supporting HCV variant replication, development of cell-free HCV replication assays, production of immunogenic HCV particles for vaccination, engineering of attenuated HCV derivatives as possible vaccine candidates, engineering of attenuated or defective HCV derivatives for expression of heterologous gene products for gene therapy and vaccine applications and for utilisation of the liver glycoproteins for targeted delivery of therapeutic agents to the liver or other cell types with appropriate receptors. Vaccine comprising these sequences is useful for inducing immunoprotection to HCV in a primate. The present sequence is Hepatitis C virus (HCV) replBartMan polyprotein variant. Note: The present sequence is not shown in the specification but is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the specification

Sequence 1985 AA;

Query Match	99.9%;	Score	10459;	DB	5;	Length	1985;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	1984;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						

QY	1	MAPITAYSQOTRGLLCIITSLTGRDNOVEGEVQVWVSTATQSFATCNGVCWTVYHGA	60
Db	1	MAPITAYSQOTRGLLCIITSLTGRDNOVEGEVQVWVSTATQSFATCNGVCWTVYHGA	60
QY	61	GSXTLAGPKGPIQMTYTNVDQDLVGWQAPPGARSLTPTCGSSDLXLVTRHADVIPVRR	120
Db	61	GSXTLAGPKGPIQMTYTNVDQDLVGWQAPPGARSLTPTCGSSDLXLVTRHADVIPVRR	120
QY	121	GDSRGSLLSPRPVSYLKGSGGELLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSEMETM	180
Db	121	GDSRGSLLSPRPVSYLKGSGGELLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSEMETM	180
QY	181	RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQYKVLVLPNSVAATLFGF	240
Db	181	RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQYKVLVLPNSVAATLFGF	240
QY	241	AYMSKAHGDIPNIRTVGRTITTCGAPITYSTYGKFLADGGCGGAYDIIICDECHSTDST	300
Db	241	AYMSKAHGDIPNIRTVGRTITTCGAPITYSTYGKFLADGGCGGAYDIIICDECHSTDST	300
QY	301	ILGIGTVLDOAETAGARLVVLATATPGSVTVPHNIEEVALSGTGEIPFYKAIPETI	360
Db	301	ILGIGTVLDOAETAGARLVVLATATPGSVTVPHNIEEVALSGTGEIPFYKAIPETI	360
QY	361	KGGRHLIFCHSKKCDLAAKLSGLGLNAVAYVYRGLDVSVIPTSGDVIVVATDLMTGFT	420
Db	361	KGGRHLIFCHSKKCDLAAKLSGLGLNAVAYVYRGLDVSVIPTSGDVIVVATDLMTGFT	420
QY	421	GDFFSVIDCMTCVTQTVDFSLDPTFTTETTPQDAVRSQRRCGTGRGMGIYRFVTPG	480



PS Claim 14; Page; 174pp; English.

XX The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) repliconMan polypeptide  
CC variant. Note: The present sequence is not shown in the specification but  
CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the  
CC specification

XX Sequence 1985 AA;

SQ Query Match 99.9%; Score 10458; DB 5; Length 1985;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLLCIIITSLTGRDRNVEGEVQVSTATQSLATCNGVCWTVYHGA 60  
DB 1 MAPITAYSQOTRGLLCIIITSLTGRDRNVEGEVQVSTATQSLATCNGVCWTVYHGA 60

QY 61 GSKTLGPKGPIOMYTNVDQLVGQAPPGASLTPTCGSDLVLTVRHADVIPVRR 120  
DB 61 GSKTLGPKGPIOMYTNVDQLVGQAPPGASLTPTCGSDLVLTVRHADVIPVRR 120

QY 121 GDSRGLLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180  
DB 121 GDSRGLLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180

QY 181 RSPVFTDNSPPAVPOTFOVAHLHAPGSGKSTKVPAAVAAQGYKVLVLPNSVAALGFG 240  
DB 181 RSPVFTDNSPPAVPOTFOVAHLHAPGSGKSTKVPAAVAAQGYKVLVLPNSVAALGFG 240

QY 241 AYMSKAGHDIPNRTGVRTITTCAPITYSYGKFLADGGCGGAYDIIICDECHSTDSTT 300  
DB 241 AYMSKAGHDIPNRTGVRTITTCAPITYSYGKFLADGGCGGAYDIIICDECHSTDSTT 300

QY 301 ILGIGTVLDAQETAGARLVVLTATPGSVTVPHNIEEVALSGTGEPFYGAIPETI 360  
DB 301 ILGIGTVLDAQETAGARLVVLTATPGSVTVPHNIEEVALSGTGEPFYGAIPETI 360

QY 361 KGRHILFCHSKKKCDELAALSGLGNNAVAYRGLDVSVIPSGDVIVVATDALMTGFT 420  
DB 361 KGRHILFCHSKKKCDELAALSGLGNNAVAYRGLDVSVIPSGDVIVVATDALMTGFT 420

QY 421 GDSVLDKNTCYTOTQVDFSLDPTFTTETTPQDVSRSORGRGTGRGMGIYRFVTPG 480  
DB 421 GDSVLDKNTCYTOTQVDFSLDPTFTTETTPQDVSRSORGRGTGRGMGIYRFVTPG 480

QY 481 ERPSGMFDSSVLCYDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWESVFTGL 540

DB 481 ERPSGMFDSSVLCYDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWESVFTGL 540

QY 541 THIDAHFLSOTKQAGDNFPYLVAYQATVCAQAAPPSPSDQMWKCLIRLKPETHGPTPLL 600  
DB 541 THIDAHFLSOTKQAGDNFPYLVAYQATVCAQAAPPSPSDQMWKCLIRLKPETHGPTPLL 600

QY 601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVSTTWLVGGVLAALAAAYCLTTSVWVIGR 660  
DB 601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVSTTWLVGGVLAALAAAYCLTTSVWVIGR 660

QY 661 IILSGKPAIPDRVLYREFDEMEECASHLPIYISQGMQABEQFKOKAIGLLQATKQAEA 720  
DB 661 IILSGKPAIPDRVLYREFDEMEECASHLPIYISQGMQABEQFKOKAIGLLQATKQAEA 720

QY 721 AAPVVEKWRLEAFWAKHWNFIISQYLAGLSTLPGNPAIASIMAFASITSLPTTQH 780  
DB 721 AAPVVEKWRLEAFWAKHWNFIISQYLAGLSTLPGNPAIASIMAFASITSLPTTQH 780

QY 781 TLLFNILGGWAAQIAPPASAASAFVAGIAGAAVSGISGLKVLDILAGYGAGVAGALVA 840  
DB 781 TLLFNILGGWAAQIAPPASAASAFVAGIAGAAVSGISGLKVLDILAGYGAGVAGALVA 840

QY 841 FKWSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPBGAGVQWNNRLIAFASR 900  
DB 841 FKWSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPBGAGVQWNNRLIAFASR 900

QY 901 GNHVSPTHVSPESDAAARVTQILSLITITQLKRLHOMINEDCSTPCSGSWLRDWDWTC 960  
DB 901 GNHVSPTHVSPESDAAARVTQILSLITITQLKRLHOMINEDCSTPCSGSWLRDWDWTC 960

QY 961 TVLTDFTKWLQSKLLPRLPGVPPFFSCQRYGKGVWGDGIMQTTCPGQAQITGHVKNXGMR 1020  
DB 961 TVLTDFTKWLQSKLLPRLPGVPPFFSCQRYGKGVWGDGIMQTTCPGQAQITGHVKNXGMR 1020

QY 1021 IVGPRTCSNTWHGTPTINAYTTGCTPSPAPNYSRALMRVAEVEVTRVGDHFYVTGM 1080  
DB 1021 IVGPRTCSNTWHGTPTINAYTTGCTPSPAPNYSRALMRVAEVEVTRVGDHFYVTGM 1080

QY 1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLREEVTVLGNLQVLVGSQLPCEPE 1140  
DB 1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLREEVTVLGNLQVLVGSQLPCEPE 1140

QY 1141 PDVAVLTSMLTDPHSHTAETAKRGLARGSPPSLAASSASQLSAPSKATCTTRHDSPDAD 1200  
DB 1141 PDVAVLTSMLTDPHSHTAETAKRGLARGSPPSLAASSASQLSAPSKATCTTRHDSPDAD 1200

QY 1201 LLEANLLMQEWMGNNITRVESKVVILDSFPLQAEDEDEREVSVAEILRRSRKFFRAM 1260  
DB 1201 LLEANLLMQEWMGNNITRVESKVVILDSFPLQAEDEDEREVSVAEILRRSRKFFRAM 1260

QY 1261 PIWARPDPNPPLESWKDPDYPPVPHGCPPLPPAKAPPIPPRRKRTVVLSESTVSALA 1320  
DB 1261 PIWARPDPNPPLESWKDPDYPPVPHGCPPLPPAKAPPIPPRRKRTVVLSESTVSALA 1320

QY 1321 ELAKTFTGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEDEPDLDGSGW 1380  
DB 1321 ELAKTFTGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEDEPDLDGSGW 1380

QY 1381 STVSEASADVCCSMYTWGTALITPCAEEETKLPINALSNLLRHNLVYATTSSAS 1440  
DB 1381 STVSEASADVCCSMYTWGTALITPCAEEETKLPINALSNLLRHNLVYATTSSAS 1440

QY 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKILLSVEEACKLTPHSARSKFGYGA 1500  
DB 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKILLSVEEACKLTPHSARSKFGYGA 1500

QY 1501 DVNLSKAVNHRSWKOLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
DB 1501 DVNLSKAVNHRSWKOLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560

QY 1561 VRVCEKMAKYDVVSTLPPQAVWGSYGFQYSGQRFVFLVNAWAKKCPMGFAYTRCFDS 1620  
DB 1561 VRVCEKMAKYDVVSTLPPQAVWGSYGFQYSGQRFVFLVNAWAKKCPMGFAYTRCFDS 1620

QY 1621 TVTENDIRVESIIYOCCLAPPEARQAIISLTERLYIGGLTNSKQNGYRRCRASGVLT 1680  
 Db 1621 TVTENDIRVESIIYOCCLAPPEARQAIISLTERLYIGGLTNSKQNGYRRCRASGVLT 1680  
 QY 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 Db 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 QY 1741 SAPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTROPTTPLAANAWEARHTFPV 1800  
 Db 1741 SAPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTROPTTPLAANAWEARHTFPV 1800  
 QY 1801 SWLGNIIWYAPTLWARMILMTHFTSILLQAQLEKALDCQYGCYSYEPDLQIQL 1860  
 Db 1801 SWLGNIIWYAPTLWARMILMTHFTSILLQAQLEKALDCQYGCYSYEPDLQIQL 1860  
 QY 1861 HGLSAPSLHSYSGEINRVASCLRLKGLVPPRLVRHRRARSVRALLSQGGRATCGKYL 1920  
 Db 1861 HGLSAPSLHSYSGEINRVASCLRLKGLVPPRLVRHRRARSVRALLSQGGRATCGKYL 1920  
 QY 1921 NWAVRTKLTPIPAASQDLSSWFVAGYSGDDIYHLSLRAPRPFWMWCLLLLSVGVGII 1980  
 Db 1921 NWAVRTKLTPIPAASQDLSSWFVAGYSGDDIYHLSLRAPRPFWMWCLLLLSVGVGII 1980  
 QY 1981 LLENR 1985  
 Db 1981 LLENR 1985  
 RESULT 12  
 ABG32460  
 ID ABG32460 standard; protein; 3010 AA.  
 AC ABG32460;  
 XX  
 DT 15-NOV-2002 (first entry)  
 DE Hepatitis C virus Con 1 isolate polyprotein mutant #9.  
 XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.  
 XX Hepatitis C virus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2199 /note= "Wild-type Ala substituted by Thr"  
 FT  
 XX WO200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX 16-JAN-2002; 2002WO-BP000526.  
 XX  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
 XX  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 XX WPI; 2002-599793/64.  
 XX  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 XX  
 XX Claim 1; Page; 69pp; English.  
 XX  
 XX The invention relates to nucleic acid molecules comprising altered HCV

CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC an HCV replicon enhanced cell or which containing a functional HCV replicon;  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
 CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
 CC and NS5B proteins), NS5A mutant of the invention. Note: The present  
 CC sequence is not shown in the specification but was created by the indexer  
 CC using the HCV sequence appearing as ABG32451 and the information in claim  
 CC 1  
 XX  
 SQ Sequence 3010 AA;

Query Match 99.9%; Score 10458; DB 5; Length 3010;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSOQTRGLGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTVHGA 60  
 Db 1026 LAPITAYSOQTRGLGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTVHGA 1085  
 QY 61 GSKTLAGPKGITOMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLXVLRHADVPVRRR 120  
 Db 1086 GSKTLAGPKGITOMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLXVLRHADVPVRRR 1145  
 QY 121 GDSRGSLLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVEMTTM 180  
 Db 1146 GDSRGSLLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVEMTTM 1205  
 QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLNPSVAATLGF 240  
 Db 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLNPSVAATLGF 1265  
 QY 241 AYMSKAHGIDPNIRTVGRTITTGAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 300  
 Db 1266 AYMSKAHGIDPNIRTVGRTITTGAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 1325  
 QY 301 ILGIGTVLDQAEATAGARLVVLATATPPGSVTVPHNIEEVALSSTGEIIFYGKAIPDIET 360  
 Db 1326 ILGIGTVLDQAEATAGARLVVLATATPPGSVTVPHNIEEVALSSTGEIIFYGKAIPDIET 1385  
 QY 361 KGRHLIFCHSKKKCDELAALKSLGLNNAVAYYRGLDYSVIFTSGDVIVVATDALMTGFT 420  
 Db 1386 KGRHLIFCHSKKKCDELAALKSLGLNNAVAYYRGLDYSVIFTSGDVIVVATDALMTGFT 1445  
 QY 421 GDFDSVIDCNTCTQTVDFSLDPTTETTTVPQDAVSRQRGRGTGRMGIIYRFTVPG 480  
 Db 1446 GDFDSVIDCNTCTQTVDFSLDPTTETTTVPQDAVSRQRGRGTGRMGIIYRFTVPG 1505  
 QY 481 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLEAYLNTPLPCVQDHLFEWESVFTGL 540  
 Db 1506 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLEAYLNTPLPCVQDHLFEWESVFTGL 1565  
 QY 541 THIDAHFLSQTQAGDNFFYLVAQATVCARAQAPPSPWDQWKKLIRLKPTLHGPTLL 600  
 Db 1566 THIDAHFLSQTQAGDNFFYLVAQATVCARAQAPPSPWDQWKKLIRLKPTLHGPTLL 1625  
 QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVI 660  
 Db 1626 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVI 1685



QY	661	IILSGKPAIIPDRVILVREFDEMEBECASHLPYIEQGMQARQFQKAKIGLLOATKQARA	720	QY	1741	SAPGDPKPKPEYDLELITSCSSNVSAHDASGKRVYVLTDRPTTPLAANAETARHTPVN	1800
Db	1686	IILSGKPAIIPDRVILVREFDEMEBECASHLPYIEQGMQARQFQKAKIGLLOATKQARA	1745	Db	2766	SAPGDPKPKPEYDLELITSCSSNVSAHDASGKRVYVLTDRPTTPLAANAETARHTPVN	2825
QY	721	AAPVVEKWRTEAPFAWKHMNFISGIQYLAGLSTLPGNPAIASMAFTASITSPITTOH	780	QY	1801	SWLGNIMYAPTLWARMLMTHFFSILLAQSEKALDCQIYGACYSIEPLDLPOLIQL	1860
Db	1746	AAPVVEKWRTEAPFAWKHMNFISGIQYLAGLSTLPGNPAIASMAFTASITSPITTOH	1805	Db	2826	SWLGNIMYAPTLWARMLMTHFFSILLAQSEKALDCQIYGACYSIEPLDLPOLIQL	2885
QY	781	TLLFNILGWWAAQAPPASAFYAGAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA	840	QY	1861	HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVHRARSVRARLLSQGGAATCGKYL	1920
Db	1806	TLLFNILGWWAAQAPPASAFYAGAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA	1865	Db	2886	HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVHRARSVRARLLSQGGAATCGKYL	2945
QY	841	PKWMSGEMPTEDLNLPAISPALVGVVCAAILRRHVGPGGAVQMMNRLIAFASR	900	QY	1921	NNAVTKLKLPIPAASQDLSSWVFAGYSGDIIYHLSRARPFRFWMCLLLSVGVGIY	1980
Db	1866	PKWMSGEMPTEDLNLPAISPALVGVVCAAILRRHVGPGGAVQMMNRLIAFASR	1925	Db	2946	NNAVTKLKLPIPAASQDLSSWVFAGYSGDIIYHLSRARPFRFWMCLLLSVGVGIY	3005
QY	901	GNHVSPTHVPESDAAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSMLRDVMDWIC	960	QY	1981	LLPNR 1985	
Db	1926	GNHVSPTHVPESDAAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSMLRDVMDWIC	1985	Db	3006	LLPNR 3010	
QY	961	TVLTDFTKWLQSKLLPRLPFGVPFFSCQYKGVWRGDMQIMOTPCGAQITGHVKNXGMR	1020	RESULT 13			
Db	1986	TVLTDFTKWLQSKLLPRLPFGVPFFSCQYKGVWRGDMQIMOTPCGAQITGHVKNXGMR	2045	ABG32454			
QY	1021	IVGPRTCSNTHGTFPINAITTCPTSPAPNYSRALMRVAEEYVEVTRVGDHVVTCM	1080	ID	ABG32454	standard; protein; 3010 AA.	
Db	2046	IVGPRTCSNTHGTFPINAITTCPTSPAPNYSRALMRVAEEYVEVTRVGDHVVTCM	2105	XX	ABG32454;		
QY	1081	TTDNVKCPQVPAPEFFTEVDGVRHLRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	1140	XX	15-NOV-2002	(first entry)	
Db	2106	TTDNVKCPQVPAPEFFTEVDGVRHLRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	2165	XX	Hepatitis C virus Con 1 isolate polyprotein mutant #3.		
QY	1141	PDVAVLTSMLTDPSSHITATKRLARGSPPSIASSASQLSAPSLKATCTTRHDSFAD	1200	XX	HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutain;		
Db	2166	PDVAVLTSMLTDPSSHITATKRLARGSPPSIASSASQLSAPSLKATCTTRHDSFAD	2225	KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;		
QY	1201	LIEANLLWQEMGNTTRVESKNVILDSFELQAEDEEREVSVAEILRSRKFPFRAM	1260	KW	internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.		
Db	2226	LIEANLLWQEMGNTTRVESKNVILDSFELQAEDEEREVSVAEILRSRKFPFRAM	2285	XX	Hepatitis C virus.		
QY	1261	PWARPDPNPPLESNKDPDYPVHVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA	1320	OS	Synthetic.		
Db	2286	PWARPDPNPPLESNKDPDYPVHVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA	2345	XX	Key	Location/Qualifiers	
QY	1321	ELATKTFGSESSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEFGDPLDSDGSW	1380	XX	Misc-difference 1347	/note= "Wild-type Ala substituted by Thr"	
Db	2346	ELATKTFGSESSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEFGDPLDSDGSW	2405	XX	WO200259321-A2.		
QY	1381	STVSEASEDVCCSMSTYWTGALITPCAEEETKLPIALNSLLRHHNLVYATTGRSAS	1440	XX	01-AUG-2002.		
Db	2406	STVSEASEDVCCSMSTYWTGALITPCAEEETKLPIALNSLLRHHNLVYATTGRSAS	2465	XX	16-JAN-2002; 2002WO-BF000526.		
QY	1441	LROKVTTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEACKLTPPHSARSKFGYGA	1500	XX	23-JAN-2001; 2001US-0263479P.		
Db	2466	LROKVTTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEACKLTPPHSARSKFGYGA	2525	XX	(RICE-) IST RICERCHE BIOL MOLECULARE ANGELETTI.		
QY	1501	DVRNLSKAVNHRSVWVKLLEDTEPIDTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560	XX	De Francesco R, Migliaccio G, Paonessa G;		
Db	2526	DVRNLSKAVNHRSVWVKLLEDTEPIDTIMAKNEVFCVQPEKGRKPARLIVFPDLG	2585	XX	WPI; 2002-599793/64.		
QY	1561	VRVCEKALYDVVSTLPAQVNGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS	1620	XX	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV		
Db	2586	VRVCEKALYDVVSTLPAQVNGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS	2645	XX	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal		
QY	1621	TVTENDIRVEESIIQCQDLAPFAQARSLRTERLYIGPLTNSKGQNCVYRCRAGVLT	1680	PT	ribosome entry site (IRES) region, useful in studying HCV replication and		
Db	2646	TVTENDIRVEESIIQCQDLAPFAQARSLRTERLYIGPLTNSKGQNCVYRCRAGVLT	2705	PT	expression.		
QY	1681	TSCGNTLTCTYKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASRAFTTEAMTRY	1740	PS	Claim 1; Page: 69pp; English.		
Db	2706	TSCGNTLTCTYKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASRAFTTEAMTRY	2765	XX	The invention relates to nucleic acid molecules comprising altered HCV		
				XX	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)		
				CC	internal ribosome entry site (IRES) region coding for one or more NS3,		
				CC	NSA, or EMCV IRES mutations, respectively. The location of the mutations		
				CC	are detailed in the specification. Also included are (1) an expression		
				CC	vector comprising a nucleotide sequence coding for the altered nucleic		
				CC	acids, which is transcriptionally coupled to an exogenous promoter; (2) a		
				CC	recombinant cell human hepatoma cell comprising the altered nucleic acids		
				CC	; (3) a recombinant cell produced by introducing into a human hepatoma		
				CC	cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)		



CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
 CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
 CC and NS5B proteins), NS3 mutant of the invention. Note: The present  
 CC sequence is not shown in the specification but was created by the indexer  
 CC using the HCV sequence appearing as ABG32451 and the information in Claim  
 CC 1  
 XX  
 SQ Sequence 3010 AA;

Query Match 99.9%; Score 10458; DB 5; Length 3010;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPITAYSOQTRGLLGCIIISLTGRDRNQVEGEVQVNSTAQSLATCVNGVCVTVYHGA 60  
 Db :  
 1026 LAPITAYSOQTRGLLGCIIISLTGRDRNQVEGEVQVNSTAQSLATCVNGVCVTVYHGA 1085

QY 61 GSKTLAGPKGPIITOMYTNVDQDLVGWQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 120  
 Db GSKTLAGPKGPIITOMYTNVDQDLVGWQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 1145

QY 121 GDSRGSLLSPRVSYLYKSGSGGPIILCPSPGHAVGIFRAAVCTRGVAXAVDFVPVSMETTM 180  
 Db GDSRGSLLSPRVSYLYKSGSGGPIILCPSPGHAVGIFRAAVCTRGVAXAVDFVPVSMETTM 1205

QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSGSKSTKVPAAAYAAQYKVLVLPNSVAATLFGF 240  
 Db RSPVFTDNSPPAVPQTFQVAHLHAPTGSGSKSTKVPAAAYAAQYKVLVLPNSVAATLFGF 1265

QY 241 AYMSKAGHDIPNRTGVRTITTCAPITYSYWYKFLADGGCGGAYDIIICDECHSTDSTT 300  
 Db AYMSKAGHDIPNRTGVRTITTCAPITYSYWYKFLADGGCGGAYDIIICDECHSTDSTT 1325

QY 301 ILGIGTVLDOAETAGARLVVLTATPPGSVTVVPHNIEEVALSGTGEIPFYGKAIPETI 360  
 Db ILGIGTVLDOAETAGARLVVLTATPPGSVTVVPHNIEEVALSGTGEIPFYGKAIPETI 1385

QY 361 KGRGHLIFCHSKKKDELAALKSLGLNAVAYYRGLDVSVIPSGDVVVVATDALMTGFT 420  
 Db KGRGHLIFCHSKKKDELAALKSLGLNAVAYYRGLDVSVIPSGDVVVVATDALMTGFT 1445

QY 421 GPDFSDVIDCMTQVTQVDFSLDPTFTIETTVPPQDAVRSORGRGTRGGMGIYRFVTPG 480  
 Db GPDFSDVIDCMTQVTQVDFSLDPTFTIETTVPPQDAVRSORGRGTRGGMGIYRFVTPG 1505

QY 481 ERPSGMFDSVLCYCEYDAGCAMELTTPAETSRLRAYLNTPLPGVQDHLFEWESVFTGL 540  
 Db ERPSGMFDSVLCYCEYDAGCAMELTTPAETSRLRAYLNTPLPGVQDHLFEWESVFTGL 1565

QY 541 THIDAHFLSQTQAGDNFYLVAQYATVCARAAPPSPDQMWKCLIRLKPTLHGPTPL 600  
 Db THIDAHFLSQTQAGDNFYLVAQYATVCARAAPPSPDQMWKCLIRLKPTLHGPTPL 1625

QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVTSTWLVGVLAALAAAYCLTTGSSVIVGR 660  
 Db YRLGAVQNEVTTTHPTTKYIMACMSADLEVTSTWLVGVLAALAAAYCLTTGSSVIVGR 1685

QY 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQAKIGLQATKQAEA 720  
 Db IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQAKIGLQATKQAEA 1745

QY 721 AAPVVESKWRITLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASMAFTASITSLTTOH 780  
 Db AAPVVESKWRITLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASMAFTASITSLTTOH 1805

QY 781 TLLFNILGGWVAQAAPPASAASFVAGIAGAAVGSIGLKVLDILAGYGAGVAGALVA 840  
 Db TLLFNILGGWVAQAAPPASAASFVAGIAGAAVGSIGLKVLDILAGYGAGVAGALVA 1865

QY 841 FKWSGEMPTSEDILVNLPAIILSPGALVGVVCAAILRRHVGPGEVQVNMRLIAFASR 900  
 Db FKWSGEMPTSEDILVNLPAIILSPGALVGVVCAAILRRHVGPGEVQVNMRLIAFASR 1925

QY 901 GNVSPHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSLRDVMDIC 960  
 Db GNVSPHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSLRDVMDIC 1985

QY 961 TVLDFDKTWLQSKLLPRLPGVPFFSCQYKGVWRGDIOMOTTCPCGAQIITHGVKNGSMR 1020  
 Db TVLDFDKTWLQSKLLPRLPGVPFFSCQYKGVWRGDIOMOTTCPCGAQIITHGVKNGSMR 2045

QY 1021 IVGPRTCSNTHWGHFTPINAYTTGCTPSPAPNYGRALWRVAAEYVEVTRVGDPHYVTGM 1080  
 Db IVGPRTCSNTHWGHFTPINAYTTGCTPSPAPNYGRALWRVAAEYVEVTRVGDPHYVTGM 2105

QY 1081 TTDNVKCPQVPAPEFFTEVDGVELHRYAPACKLLREEVTFVLGVLNQYLVGSLPCEPE 1140  
 Db TTDNVKCPQVPAPEFFTEVDGVELHRYAPACKLLREEVTFVLGVLNQYLVGSLPCEPE 2165

QY 1141 PDVAVLTSMLTDPESHITAEAKRLARSGPPSLASSASQSLAPSLKATCTTRHDSPAD 1200  
 Db PDVAVLTSMLTDPESHITAEAKRLARSGPPSLASSASQSLAPSLKATCTTRHDSPAD 2225

QY 1201 LIEANLLWROEMGNIITRVSEKNIWILDSFEPLOADEDEREVSVPABILRRSRKPPRAM 1260  
 Db LIEANLLWROEMGNIITRVSEKNIWILDSFEPLOADEDEREVSVPABILRRSRKPPRAM 2285

QY 1261 PIWARPDPYNPLLESWKDPDYVPVHVHGCPLPAKAPPIPPBRKRVTWLSESTVSALA 1320  
 Db PIWARPDPYNPLLESWKDPDYVPVHVHGCPLPAKAPPIPPBRKRVTWLSESTVSALA 2345

QY 1321 ELATKTFGSESSAIVSGTATASPDQSPDDGAGSDVESYSSMPPLGEFGDPLDSGSW 1380  
 Db ELATKTFGSESSAIVSGTATASPDQSPDDGAGSDVESYSSMPPLGEFGDPLDSGSW 2405

QY 1381 STVSEESAEEDVCCSYTWTGALITPCAABETKLPINALNSLLRHNNLVYATTSRAS 1440  
 Db STVSEESAEEDVCCSYTWTGALITPCAABETKLPINALNSLLRHNNLVYATTSRAS 2465

QY 1441 LRQKVTFRQLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK 1500  
 Db LRQKVTFRQLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK 2525

QY 1501 DVRLSSKAVNHIRSVMKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 Db DVRLSSKAVNHIRSVMKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585

QY 1561 VRVCEKALYDVVSTLPQAVMGSSYGFQSPQORVEFLVNAWKAKCPMGFAVDTTCFDS 1620  
 Db VRVCEKALYDVVSTLPQAVMGSSYGFQSPQORVEFLVNAWKAKCPMGFAVDTTCFDS 2645

QY 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGNCCYVRCRASGVLT 1680  
 Db TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGNCCYVRCRASGVLT 2705

QY 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVIICESAGTOEDEASLRAFTAMTRY 1740  
 Db TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVIICESAGTOEDEASLRAFTAMTRY 2765

QY 1741 SAPPDGPPEYDLELITSCSNNSVAHDASGRVYLYLTRDPTTPLARAAMETARTHPN 1800  
 Db SAPPDGPPEYDLELITSCSNNSVAHDASGRVYLYLTRDPTTPLARAAMETARTHPN 2825

QY 1801 SWLGNIMYAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPDLPOIQL 1860  
 Db SWLGNIMYAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPDLPOIQL 2885

QY 1861 HGLSAFSLHSYSPGEBINRVASCLRKLGVPLURVWRHRARSVRARLLSOGGRAATCGKYL 1920

Db 2886 HGLSAFSLHSYSPGEINRVASCLRLKLGVPURVWRHRSVRARLLSOGGRRATCGKLYF 2945  
 QY 1921 NWAVRTKLTPIPAASOLDLSSWFVAGYSGGDIYHSLSRAPRPFMMWCLLLLSVGVGIY 1980  
 Db 2946 NWAVRTKLTPIPAASOLDLSSWFVAGYSGGDIYHSLSRAPRPFMMWCLLLLSVGVGIY 3005

QY 1981 LLPNR 1985  
 Db 3006 LLPNR 3010

RESULT 14  
 ABG32461  
 ID ABG32461 standard; protein; 3010 AA.

AC ABG32461;

DT 15-NOV-2002 (first entry)

DE Hepatitis C virus Con 1 isolate polyprotein mutant #10.

KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten; HCV;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.

OS Hepatitis C virus.  
 OS Synthetic.

FX Key Location/Qualifiers

FT Misc-difference 2204 /note= "wild-type Ser substituted by Arg"

FT WO200259321-A2.

FX 01-AUG-2002.

FX 16-JAN-2002; 2002WO-EP000526.

FX 23-JAN-2001; 2001US-0263479P.

FX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

FX De Francesco R, Migliaccio G, Paonessa G;

FX MPI; 2002-599793/64.

FX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 ribosome entry site (IRES) region, useful in studying HCV replication and  
 expression.

FX Claim 1; Page; 69pp; English.

FX The invention relates to nucleic acid molecules comprising altered HCV  
 NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 internal ribosome entry site (IRES) region coding for one or more NS3,  
 NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 are detailed in the specification. Also included are (1) an expression  
 vector comprising a nucleotide sequence coding for the altered nucleic  
 acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 recombinant cell human hepatoma cell comprising the altered nucleic acids  
 ; (3) a recombinant cell produced by introducing into a human hepatoma  
 cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 replicon enhanced cell or which containing a functional HCV replicon; (5)  
 an HCV replicon enhanced cells made in the method; and (6) measuring the  
 ability of a compound to affect HCV activity. The HCV replicons and HCV  
 replicon enhanced cells are useful in studying HCV replication and  
 expression, and HCV and host cell interactions, producing HCV RNA and  
 proteins, and providing a system for measuring the ability of a compound  
 to modulate one or more HCV activities e.g. to discover drugs which may  
 treat HCV mediated diseases such as liver failure, cirrhosis and  
 hepatocellular carcinoma. The present sequence is the HCV replicon Con 1

CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
 CC and NS5B proteins). NS5A mutant of the invention. Note: The present  
 CC sequence is not shown in the specification but was created by the indexer  
 CC using the HCV sequence appearing as ABG32451 and the information in claim  
 CC 1  
 XX SQ Sequence 3010 AA;

Query Match 99.9%; Score 10457; DB 5; Length 3010;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPTAYSQQRGLGCIITSITGRDRNQVGEVQVSTATQSFATCNGVCWTVYHGA 60  
 Db 1026 LAPITAYSQQRGLGCIITSITGRDRNQVGEVQVSTATQSFATCNGVCWTVYHGA 1085  
 QY 61 GSKTLAGPKGPITQMYTNVDQDLVGMQAPPGARSITPCTCGSSDLYLVTRHADVLPVRRR 120  
 Db 1086 GSKTLAGPKGPITQMYTNVDQDLVGMQAPPGARSITPCTCGSSDLYLVTRHADVLPVRRR 1145  
 QY 121 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
 Db 1146 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
 QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKSTKYPAAVAGYKVLVNPVAATLGG 240  
 Db 1206 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKSTKYPAAVAGYKVLVNPVAATLGG 1265  
 QY 241 AYMSKAHGIDENIRTCVRTITTTGAPITYSTYKFLADGCGSGAGYDIIICDECHSDSTT 300  
 Db 1266 AYMSKAHGIDENIRTCVRTITTTGAPITYSTYKFLADGCGSGAGYDIIICDECHSDSTT 1325  
 QY 301 ILGIGTVLDQARTAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYGKAIPETI 360  
 Db 1326 ILGIGTVLDQARTAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYGKAIPETI 1385  
 QY 361 KGRHLIFCHSKKKCDELAALKSLGLNNAVAYRGDLVSIVTSGDVIIVADALMTGT 420  
 Db 1386 KGRHLIFCHSKKKCDELAALKSLGLNNAVAYRGDLVSIVTSGDVIIVADALMTGT 1445  
 QY 421 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTVQDAVSRSQRRGTRGGRMGYRVTTPG 480  
 Db 1446 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTVQDAVSRSQRRGTRGGRMGYRVTTPG 1505  
 QY 481 ERPSGMFSDSVLCECYDAGCAWYELTPAETSRLRAYLNTPLGPVCDHLFPWESVFTGL 540  
 Db 1506 ERPSGMFSDSVLCECYDAGCAWYELTPAETSRLRAYLNTPLGPVCDHLFPWESVFTGL 1565  
 QY 541 THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDQWKCLIRLKPTLHGPTPL 600  
 Db 1566 THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDQWKCLIRLKPTLHGPTPL 1625  
 QY 601 YRLGAVONEVTTTHPIITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
 Db 1626 YRLGAVONEVTTTHPIITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 1685  
 QY 661 IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLQTATQAE 720  
 Db 1686 IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLQTATQAE 1745  
 QY 721 AAPVSKMTLEAFWAKHWNFIISGIQYLAGISTLPGNPAIASLMFTAFTASITSLTQ 780  
 Db 1746 AAPVSKMTLEAFWAKHWNFIISGIQYLAGISTLPGNPAIASLMFTAFTASITSLTQ 1805  
 QY 781 TLLFNILGGVAAQLAPPSSAASAFVAGIAGAAGVSTGLGKLVLDILAGYAGAGVAGALVA 840  
 Db 1806 TLLFNILGGVAAQLAPPSSAASAFVAGIAGAAGVSTGLGKLVLDILAGYAGAGVAGALVA 1865  
 QY 841 FKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILIAFASR 900  
 Db 1866 FKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILIAFASR 1925  
 QY 901 GNHVSPTHYVPESDAAARVQTILSSLLITQLLKHQLWINEDECSTPCSGSWLRDWDWIC 960

Db 1926 GNVSTHVPESDAARVTOILSSITITQLLKRLHQWINEDESTPCSGSWLRDWDWIC 1985  
QY 961 TVLTDFKTLQSKLLPRLPGVPFFSCQGYKGYWVGDMQITTCPCGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFKTLQSKLLPRLPGVPFFSCQGYKGYWVGDMQITTCPCGAQITGHVKNQSMR 2045  
QY 1021 INVPRTCSTWGTTFPINAYTTGPTCPSPAPNYSRALNRVAABEYVEVTRVGDHYVTGM 1080  
Db 2046 INVPRTCSTWGTTFPINAYTTGPTCPSPAPNYSRALNRVAABEYVEVTRVGDHYVTGM 2105  
QY 1081 TTDNVKPCQVPAPEPTEVDGVRHRYAPACKPLLEEVTVLVGLNQYLVGSQLCEPE 1140  
Db 2106 TTDNVKPCQVPAPEPTEVDGVRHRYAPACKPLLEEVTVLVGLNQYLVGSQLCEPE 2165  
QY 1141 PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLIASSASQLSAPSLKATCTTRHDSPAD 1200  
Db 2166 PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLIASSASQLSAPSLKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWQEMGNLTRVESNKVILDSFELQAEEDEREVSVAEILRRSRKFPFRAM 1260  
Db 2226 LIEANLLWQEMGNLTRVESNKVILDSFELQAEEDEREVSVAEILRRSRKFPFRAM 2285  
QY 1261 PIWAPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 2286 PIWAPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 2345  
QY 1321 ELATKTFGSSSASVDSGTATASPDQSDGDAGSDVESYSPMPLEGPDPDLSDGSW 1380  
Db 2346 ELATKTFGSSSASVDSGTATASPDQSDGDAGSDVESYSPMPLEGPDPDLSDGSW 2405  
QY 1381 STVSEASDVVCCSMYSYTWGALITPCAETKLPINALSNLSLHNLVYVATTSRAS 1440  
Db 2406 STVSEASDVVCCSMYSYTWGALITPCAETKLPINALSNLSLHNLVYVATTSRAS 2465  
QY 1441 LRQKVTFRDLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
Db 2466 LRQKVTFRDLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 2525  
QY 1501 DVYRNSSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPALIVFPDYG 1560  
Db 2526 DVYRNSSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPALIVFPDYG 2585  
QY 1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVBEFLVNAWKAKKCPMGFYDTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVBEFLVNAWKAKKCPMGFYDTRCFDS 2645  
QY 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGLPTNSKGONCCGYRRCRAGSVLT 1680  
Db 2646 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGLPTNSKGONCCGYRRCRAGSVLT 2705  
QY 1681 TSCGNLTLCYLAACAAKRAKIQDCTMLVCGDDLVVICESAGTQBEASLRAFTAMTRY 1740  
Db 2706 TSCGNLTLCYLAACAAKRAKIQDCTMLVCGDDLVVICESAGTQBEASLRAFTAMTRY 2765  
QY 1741 SAPPDGPPEYDLELITSCSNVSVAHDASGRVYVLTTRDPTPLARAAMTARHTPN 1800  
Db 2766 SAPPDGPPEYDLELITSCSNVSVAHDASGRVYVLTTRDPTPLARAAMTARHTPN 2825  
QY 1801 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIOL 1860  
Db 2826 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIOL 2885  
QY 1861 HGLSAPSLHSYSGEINRVASCLRLGVPLPLVRHRRARSVRARLLSQGGRAATCGKYL 1920  
Db 2886 HGLSAPSLHSYSGEINRVASCLRLGVPLPLVRHRRARSVRARLLSQGGRAATCGKYL 2945  
QY 1921 NNAVTKLTLTIPAAQQLDLSWVAGYSGGDIYHLSRARPRFWMWCLLLSVGVGIY 1980  
Db 2946 NNAVTKLTLTIPAAQQLDLSWVAGYSGGDIYHLSRARPRFWMWCLLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
|||||

Db 3006 LLPNR 3010  
RESULT 15  
ID ABG32458  
XX ABG32458 standard; protein; 3010 AA.  
XX AC ABG32458;  
XX DT 15-NOV-2002 (first entry)  
XX Hepatitis C virus Con 1 isolate polyprotein mutant #7.  
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutain;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.  
XX Hepatitis C virus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 2197  
FT /note= "Wild-type Phe substituted by Ser"  
XX WO200259321-A2.  
XX 01-AUG-2002.  
XX 16-JAN-2002; 2002WO-BP000526.  
XX 23-JAN-2003; 2001US-0263479P.  
XX (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX WPI; 2002-599793/64.  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX Claim 1; Page; 69pp; English.  
XX  
CC The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
CC and NS5B proteins), NS5A mutant of the invention. Note: The present  
CC sequence is not shown in the specification but was created by the indexer  
CC using the HCV sequence appearing as ABG32451 and the information in claim  
CC 1  
XX Sequence 3010 AA;  
SQ

Query Match 99.9%; Score 10456; DB 5; Length 3010;

Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	MAPTAYSQQRGLGCIITSLTGRDRNQVEGVVSTATOSFLATCVNGVCWTVYHGA	60
Db	1026	LAPITAYSQQRGLGCIITSLTGRDRNQVEGVVSTATOSFLATCVNGVCWTVYHGA	1085
Qy	61	GSKTLAGPKGPI TOMYTNVDQDLVQWQAPPGARSITPCTCGSSDLYL VTRHADVIPVRR	120
Db	1086	GSKTLAGPKGPI TOMYTNVDQDLVQWQAPPGARSITPCTCGSSDLYL VTRHADVIPVRR	1145
Qy	121	GDSRGLSPRPVSYLKGSSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVSMETTM	180
Db	1146	GDSRGLSPRPVSYLKGSSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVSMETTM	1205
Qy	181	RSPVFTDNSSPPAVPQTQVAHLHAPTGSKSTKVPAAYAAQGYKVLVNLPSVAATLFG	240
Db	1206	RSPVFTDNSSPPAVPQTQVAHLHAPTGSKSTKVPAAYAAQGYKVLVNLPSVAATLFG	1265
Qy	241	AYMSKAHGIDNIRTVGTITGTAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	300
Db	1266	AYMSKAHGIDNIRTVGTITGTAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	1325
Qy	301	ILGIGTVLDQAETAGARLVWLATATPPGSVTVPHENIEEVALSSSTGEIPFYGKAIPETI	360
Db	1326	ILGIGTVLDQAETAGARLVWLATATPPGSVTVPHENIEEVALSSSTGEIPFYGKAIPETI	1385
Qy	361	KGGRHLIFCHSKKKDELAALKSLGLGNAYYRGLDVSVIPTSDDVIVVATDALMTGFT	420
Db	1386	KGGRHLIFCHSKKKDELAALKSLGLGNAYYRGLDVSVIPTSDDVIVVATDALMTGFT	1445
Qy	421	GDFSVIDCNTCVTVDFSLDPTTETITTVQDAVSRSORGRGTGRMGIIYRFVTPG	480
Db	1446	GDFSVIDCNTCVTVDFSLDPTTETITTVQDAVSRSORGRGTGRMGIIYRFVTPG	1505
Qy	481	ERPSCMFDSVLCYCEYDAGCWAYELTPAETS VRLRAYLNTPLGVPQCDELFEFWSVFTGL	540
Db	1506	ERPSCMFDSVLCYCEYDAGCWAYELTPAETS VRLRAYLNTPLGVPQCDELFEFWSVFTGL	1565
Qy	541	THIDAHFLSQTQAGDNPPYL VAYQATVCARAQAPPPSDQWQKCLIRL KPTLHPTPLL	600
Db	1566	THIDAHFLSQTQAGDNPPYL VAYQATVCARAQAPPPSDQWQKCLIRL KPTLHPTPLL	1625
Qy	601	YRLGAVQNEVTTTHPIKYIMACMSADLEVNTSTWLVGGVLAALAAVCLITGVSVVIGR	660
Db	1626	YRLGAVQNEVTTTHPIKYIMACMSADLEVNTSTWLVGGVLAALAAVCLITGVSVVIGR	1685
Qy	661	IILSGKPAIIPREVLYREFDEMEBCASHLPYIEQGMQLAEQFKQKAILGLLTATKQAEA	720
Db	1686	IILSGKPAIIPREVLYREFDEMEBCASHLPYIEQGMQLAEQFKQKAILGLLTATKQAEA	1745
Qy	721	AAPVYESKWRTEAFWAKHMNFISGIQYLAGLSTLPGNPATASLMAFTASITSLPTTQH	780
Db	1746	AAPVYESKWRTEAFWAKHMNFISGIQYLAGLSTLPGNPATASLMAFTASITSLPTTQH	1805
Qy	781	TLLFNILGWAQAOLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVA	840
Db	1806	TLLFNILGWAQAOLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVA	1865
Qy	841	FKVMSGEMPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVQWMMNRLIAFASR	900
Db	1866	FKVMSGEMPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVQWMMNRLIAFASR	1925
Qy	901	GNHVSPTHVPESDAAARTQILSLTITQLLKRHLQWINEDCSTPCSGSWLRDWDWIC	960
Db	1926	GNHVSPTHVPESDAAARTQILSLTITQLLKRHLQWINEDCSTPCSGSWLRDWDWIC	1985
Qy	961	TVLTDFTKWLQSKLLPRLPVGFVFFSCQRYKGVWRGDI MQTTCPGQAQITGHVXNGSMR	1020
Db	1986	TVLTDFTKWLQSKLLPRLPVGFVFFSCQRYKGVWRGDI MQTTCPGQAQITGHVXNGSMR	2045
Qy	1021	IVGPRTCSNTWHGTTPINAYTTGPTSPAPNYSRALWRVAABEEVYVTRVGDFFHYVTGM	1080

Db	2046	IVGPRTCSNTWHGTTPINAYTTGPTSPAPNYSRALWRVAABEEVYVTRVGDFFHYVTGM	2105
Qy	1081	TTDNVVKPCQVPAPBEFFTEVDGVPLHRYAPACKPLIREEVTFLYGLNOLYVCSQLPCPE	1140
Db	2106	TTDNVVKPCQVPAPBEFFTEVDGVPLHRYAPACKPLIREEVTFLYGLNOLYVCSQLPCPE	2165
Qy	1141	PDVAVLTSMLTDPGSHITAEAKRRLARGSPPLASSASQLSAPSLKATCTTRHDSPAD	1200
Db	2166	PDVAVLTSMLTDPGSHITAEAKRRLARGSPPLASSASQLSAPSLKATCTTRHDSPAD	2225
Qy	1201	LIEANLWRQBMGNITRVESENKVIILDSPEPIQAEEDEREVSVPABILRRSRKFPFRAM	1260
Db	2226	LIEANLWRQBMGNITRVESENKVIILDSPEPIQAEEDEREVSVPABILRRSRKFPFRAM	2285
Qy	1261	PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTTVLSESTVSSALA	1320
Db	2286	PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTTVLSESTVSSALA	2345
Qy	1321	ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSMPPLGEPEGDPDLSDGSM	1380
Db	2346	ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSMPPLGEPEGDPDLSDGSM	2405
Qy	1381	STVSEASEDVVCCSMSTYWTGALITPCAABETKLPINALNSLLRHHNLVYATTSRAS	1440
Db	2406	STVSEASEDVVCCSMSTYWTGALITPCAABETKLPINALNSLLRHHNLVYATTSRAS	2465
Qy	1441	LROKKTVPDRLOVLDHVDYLVKEMKAKASTVKAKLSVEBACKLTPPHSARSKEGYAK	1500
Db	2466	LROKKTVPDRLOVLDHVDYLVKEMKAKASTVKAKLSVEBACKLTPPHSARSKEGYAK	2525
Qy	1501	DVRNLSSKAVNHRISVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPARLI VFPDLG	1560
Db	2526	DVRNLSSKAVNHRISVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPARLI VFPDLG	2585
Qy	1561	VRVCEKMALYDVVSTLPOAVMGSSYGFQYSGQVFEFLVNAWKAKKCPMGFAYDTRCFDS	1620
Db	2586	VRVCEKMALYDVVSTLPOAVMGSSYGFQYSGQVFEFLVNAWKAKKCPMGFAYDTRCFDS	2645
Qy	1621	TVTENDIRVESIYOCCLDAPARQAIRSLTERLYIGGLPLTNSKGQNGYRRCRASGVLT	1680
Db	2646	TVTENDIRVESIYOCCLDAPARQAIRSLTERLYIGGLPLTNSKGQNGYRRCRASGVLT	2705
Qy	1681	TSCGNLTLCYLKAAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEASLRAFTAMTRY	1740
Db	2706	TSCGNLTLCYLKAAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEASLRAFTAMTRY	2765
Qy	1741	SAPPGDPKPEYDLELITSCSSNSVSVAHDAAGKRVYVITRDPPTPLAAMETARHTPVN	1800
Db	2766	SAPPGDPKPEYDLELITSCSSNSVSVAHDAAGKRVYVITRDPPTPLAAMETARHTPVN	2825
Qy	1801	SWLGNIIMVAPTLWARMILMTHFFSILLAOEKLALDCQIYGACYSIEPLDLPOLIORL	1860
Db	2826	SWLGNIIMVAPTLWARMILMTHFFSILLAOEKLALDCQIYGACYSIEPLDLPOLIORL	2885
Qy	1861	HGLSAFSLHSYSPGEBINRVASCLRLGVPPPLRVWHRARSVRARLLSOGGAAATCGKYL	1920
Db	2886	HGLSAFSLHSYSPGEBINRVASCLRLGVPPPLRVWHRARSVRARLLSOGGAAATCGKYL	2945
Qy	1921	NWAVTKLKLTPIPAASQDLSSWFVAGYSGGDIYHLSRARPRFMKCLLLLSVGVGLY	1980
Db	2946	NWAVTKLKLTPIPAASQDLSSWFVAGYSGGDIYHLSRARPRFMKCLLLLSVGVGLY	3005
Qy	1981	LLPNR 1985	
Db	3006	LLPNR 3010	

Search completed: December 8, 2004, 12:27:28  
Job time : 147 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:20:57 ; Search time 33 Seconds  
(without alignments)  
3989.128 Million cell updates/sec

Title: US-09-576-989-3  
Perfect score: 10465  
Sequence: 1 MAPITAYSQOTRGLGCIIT.....FMWCLLLLSVGVGVYLLPNR 1985

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10465	100.0	1985	4	US-09-539-601-9
2	10465	100.0	1985	4	US-09-539-601-12
3	10462	100.0	2201	4	US-09-539-601-6
4	10462	100.0	2201	4	US-09-539-601-15
5	10462	100.0	3010	4	US-09-539-601-3
6	10442	99.8	2201	4	US-10-029-907-3
7	10432	99.7	1985	4	US-09-539-601-24
8	10432	99.7	3010	4	US-09-539-601-27
9	10425	99.6	1985	4	US-09-539-601-18
10	10422	99.6	3010	4	US-09-539-601-21
11	10412	99.5	1985	4	US-09-539-601-30
12	10409	99.5	3010	4	US-09-539-601-33
13	10170	97.2	2201	3	US-08-952-981A-2
14	10170	97.2	2621	2	US-08-324-977-36
15	10170	97.2	2621	2	US-08-384-616-36
16	10170	97.2	2621	2	US-08-304-686A-36
17	10170	97.2	2621	3	US-09-315-850-36
18	10170	97.2	3010	1	US-08-324-977-2
19	10170	97.2	3010	1	US-08-324-977-14
20	10170	97.2	3010	2	US-08-384-616-2
21	10170	97.2	3010	2	US-08-384-616-14
22	10170	97.2	3010	2	US-08-304-686A-2
23	10170	97.2	3010	2	US-08-304-686A-14
24	10170	97.2	3010	3	US-09-315-850-2
25	10170	97.2	3010	3	US-09-315-850-14
26	10165	97.1	2620	1	US-08-324-977-32
27	10165	97.1	2620	2	US-08-384-616-32

28	10165	97.1	2620	2	US-08-904-686A-32	Sequence 32, Appl
29	10165	97.1	2620	3	US-09-315-850-32	Sequence 32, Appl
30	10106	96.6	3010	3	US-09-014-416-3	Sequence 3, Appl
31	9430.5	90.1	3012	3	US-08-811-566-2	Sequence 2, Appl
32	9430.5	90.1	3012	3	US-09-034-756-2	Sequence 2, Appl
33	9426.5	90.1	3011	1	US-08-453-552-2	Sequence 2, Appl
34	9426.5	90.1	3011	2	US-08-710-637-2	Sequence 2, Appl
35	9426.5	90.1	3011	3	US-08-650-328-1	Sequence 1, Appl
36	9426.5	90.1	3011	5	PCT-US93-00907-2	Sequence 1, Appl
37	9418.5	90.0	3011	3	US-09-014-416-1	Sequence 5, Appl
38	9418.5	90.0	3011	4	US-09-952-572-9	Sequence 9, Appl
39	9417.5	90.0	3011	1	US-08-440-103-36	Sequence 36, Appl
40	9417.5	90.0	3011	1	US-08-440-542-36	Sequence 36, Appl
41	9417.5	90.0	3011	1	US-07-910-760-10	Sequence 10, Appl
42	9417.5	90.0	3011	1	US-08-440-519-10	Sequence 10, Appl
43	9417.5	90.0	3011	1	US-08-231-368-36	Sequence 36, Appl
44	9417.5	90.0	3011	1	US-08-440-210-36	Sequence 36, Appl
45	9417.5	90.0	3011	1		

ALIGNMENTS

RESULT 1  
US-09-539-601-9  
; Sequence 9, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartschlag, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-9

Query Match	100.0%;	Score	10465;	DB	4;	Length	1985;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1985;	Conservative	0;	Mismatches	0;	Indels	0;
		Gaps	0;				
QY	1	MAPITAYSQOTRGLGCIIT	SLTGRDRNQVEGVVSTATOSFLATCVNGVCTVYHGA	60			
Db	1	MAPITAYSQOTRGLGCIIT	SLTGRDRNQVEGVVSTATOSFLATCVNGVCTVYHGA	60			
QY	61	GSKTLAGPKGPITQMTNVNDQ	LVGWQAPPGARSLTPTCTCGSSDLVLTTHADVIPVRR	120			
Db	61	GSKTLAGPKGPITQMTNVNDQ	LVGWQAPPGARSLTPTCTCGSSDLVLTTHADVIPVRR	120			
QY	121	GSRSGLSPRVSVLYKSGSGPL	ICPSGHAVIFRAAVCTRGVAKAVDFVPVSEMETM	180			
Db	121	GSRSGLSPRVSVLYKSGSGPL	ICPSGHAVIFRAAVCTRGVAKAVDFVPVSEMETM	180			
QY	181	RSPVFTDNSSPAVPTQFOVHL	HAPTGSKSTKVPAAVAAOGYKVLNPSVAATLGF	240			
Db	181	RSPVFTDNSSPAVPTQFOVHL	HAPTGSKSTKVPAAVAAOGYKVLNPSVAATLGF	240			
QY	241	AYWSXAHGIDPNIRTVGRTIT	TGAPITTYSTYKFLADGGCGGAYDIIICDECHSDSTT	300			
Db	241	AYWSXAHGIDPNIRTVGRTIT	TGAPITTYSTYKFLADGGCGGAYDIIICDECHSDSTT	300			
QY	301	ILGIGTVLDQATAGARLVWL	TATPPGSTVTPHPNIEVALSSGTGEIPFYKAPIETI	360			
Db	301	ILGIGTVLDQATAGARLVWL	TATPPGSTVTPHPNIEVALSSGTGEIPFYKAPIETI	360			
QY	361	KGGRHLIFCHSKKKDELA	AKLSGLNAVAYRGDUSVIETSGDIVVATDALMTGT	420			

Db 361 KGRHLIFCHSKKKDELAALKSLGLGLNNAVAYRGLDVSVIPTSGDVIWATDALMTGFT 420  
Qy 421 GDFSVDCNTCVCTQVDFSLDPTETIETTTVPQDAVRSORRGHTGRMGIRYFVTPG 480  
Db 421 GDFSVDCNTCVCTQVDFSLDPTETIETTTVPQDAVRSORRGHTGRMGIRYFVTPG 480  
Qy 481 ERPSGMFSSVLCYCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQDHLFEWESVFTGL 540  
Db 481 ERPSGMFSSVLCYCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQDHLFEWESVFTGL 540  
Qy 541 THIDAHFTSOTKQAGDNPPYLVAYQATVCARAQAPPPSDQWQKCLIRKPTLHGPTPLL 600  
Db 541 THIDAHFTSOTKQAGDNPPYLVAYQATVCARAQAPPPSDQWQKCLIRKPTLHGPTPLL 600  
Qy 601 YRLGAVQNEVTTTHPIIKYINMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIYGR 660  
Db 601 YRLGAVQNEVTTTHPIIKYINMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIYGR 660  
Qy 661 IILSGKPAIIPDREVLYREFDEMBCASHLPYIEQGMQLAEQFKOKAIGLLQTATKQABA 720  
Db 661 IILSGKPAIIPDREVLYREFDEMBCASHLPYIEQGMQLAEQFKOKAIGLLQTATKQABA 720  
Qy 721 AAPVVESKWRTELEAFWAKHMFNFGIYQIYLAGSLTPGNPAIASLMAFTASITSLTTOH 780  
Db 721 AAPVVESKWRTELEAFWAKHMFNFGIYQIYLAGSLTPGNPAIASLMAFTASITSLTTOH 780  
Qy 781 TLLFNILGGWAAQALAPPSAASAFVAGAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 840  
Db 781 TLLFNILGGWAAQALAPPSAASAFVAGAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 840  
Qy 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGGEGAVQWNNRLJAFASR 900  
Db 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGGEGAVQWNNRLJAFASR 900  
Qy 901 GNVHSPHYVPESDAARVTOILSLTITQLLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960  
Db 901 GNVHSPHYVPESDAARVTOILSLTITQLLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960  
Qy 961 TVLDFKTLWOSKLLPLPGVPFFSCORGKGVWGGIMQITPCGAQITGHVKNKSGMR 1020  
Db 961 TVLDFKTLWOSKLLPLPGVPFFSCORGKGVWGGIMQITPCGAQITGHVKNKSGMR 1020  
Qy 1021 IVGPRTCNTWHGTFPINAYTGTCTSPAPNYSRALMRVAEEYVEVTRVCGDFHYVTGM 1080  
Db 1021 IVGPRTCNTWHGTFPINAYTGTCTSPAPNYSRALMRVAEEYVEVTRVCGDFHYVTGM 1080  
Qy 1081 TTDNVKCPQVAPBEFFTEVDGVRHLRYAPACKPILLREEVFLVGLNQYLVGSQLPCEPPE 1140  
Db 1081 TTDNVKCPQVAPBEFFTEVDGVRHLRYAPACKPILLREEVFLVGLNQYLVGSQLPCEPPE 1140  
Qy 1141 PDVAVLTSMITDPSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Db 1141 PDVAVLTSMITDPSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Qy 1201 LIEANLLWRQEMGNITRVESENKVILDSFEPLQAEEDEREVSVPAILRRSRKFFPRAM 1260  
Db 1201 LIEANLLWRQEMGNITRVESENKVILDSFEPLQAEEDEREVSVPAILRRSRKFFPRAM 1260  
Qy 1261 PIWARPDPNPLLESKWDPDVVPVHGCPIPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1261 PIWARPDPNPLLESKWDPDVVPVHGCPIPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Qy 1321 ELATKTFGSSSSAVDSGTATASDPQSDDDGAGSDVESYSMPLEBGPDPDLSGWSW 1380  
Db 1321 ELATKTFGSSSSAVDSGTATASDPQSDDDGAGSDVESYSMPLEBGPDPDLSGWSW 1380  
Qy 1381 STVSEASEDVVCCSMSTWTGALITPCAAEBETKLPIINALNSLRLHNLVYATTSRSAS 1440  
Db 1381 STVSEASEDVVCCSMSTWTGALITPCAAEBETKLPIINALNSLRLHNLVYATTSRSAS 1440  
Qy 1441 LROKVVTFDRLOVLDHDDYDLVKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK 1500  
Db 1441 LROKVVTFDRLOVLDHDDYDLVKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK 1500

Qy 1501 DVRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGGKRPARLIIVFPDLG 1560  
Db 1501 DVRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGGKRPARLIIVFPDLG 1560  
Qy 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQGRVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Db 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQGRVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Qy 1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKQNGCYRRCRASGVLT 1680  
Db 1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKQNGCYRRCRASGVLT 1680  
Qy 1681 TSCGNLTLCYLAACAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTAMTRY 1740  
Db 1681 TSCGNLTLCYLAACAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTAMTRY 1740  
Qy 1741 SAPPGDPKPYDLELITSCSSNSVSAHDASGKRVYVYLRDPTTPLAARAAWETAHTPTVN 1800  
Db 1741 SAPPGDPKPYDLELITSCSSNSVSAHDASGKRVYVYLRDPTTPLAARAAWETAHTPTVN 1800  
Qy 1801 SWLGNIIWYATLWARMLIMTHFFSILLAOBLEKALDCQIYGACYSIEPDLPOIIOQL 1860  
Db 1801 SWLGNIIWYATLWARMLIMTHFFSILLAOBLEKALDCQIYGACYSIEPDLPOIIOQL 1860  
Qy 1861 HGLSAPSLHSYSPGEINRVASCLRKLGVPPLRVWRHRSVRARLLSOGGRAATCGKYL 1920  
Db 1861 HGLSAPSLHSYSPGEINRVASCLRKLGVPPLRVWRHRSVRARLLSOGGRAATCGKYL 1920  
Qy 1921 NNAVRTKLKLTPIPAASQDLSSMFVAGYSGGDIYHLSLRARPRFWMCLLLSSVGVGIY 1980  
Db 1921 NNAVRTKLKLTPIPAASQDLSSMFVAGYSGGDIYHLSLRARPRFWMCLLLSSVGVGIY 1980  
Qy 1981 LLPNR 1985  
Db 1981 LLPNR 1985

## RESULT 2

US-09-539-601-12  
; Sequence 12, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-12

Query Match 100.0%; Score 10465; DB 4; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPITATYSQQTRGLLGCIIITSLTGRDRNQVGEVQVWSTATQSFATCVNGVCWTVYHGA 60  
Db 1 MAPITATYSQQTRGLLGCIIITSLTGRDRNQVGEVQVWSTATQSFATCVNGVCWTVYHGA 60  
Qy 61 GSKTLAGPKGPIQWYTNVDQDLVQWQAPPQARSITPTCTCGSSDLYLVRHADVIPVRRR 120  
Db 61 GSKTLAGPKGPIQWYTNVDQDLVQWQAPPQARSITPTCTCGSSDLYLVRHADVIPVRRR 120  
Qy 121 GDSRGSLLSPRPVSVLYKSGSGGLLCPSHAVGIFPRAAVCTRGVAKAVDFEVPESMETTM 180  
Db 121 GDSRGSLLSPRPVSVLYKSGSGGLLCPSHAVGIFPRAAVCTRGVAKAVDFEVPESMETTM 180

181 RSPVFTDNSSPPAVPOTFOVAHLHAPITGSGSKTKVPAAYAAQGYKVLVLNPSVAATLGF 240  
 181 RSPVFTDNSSPPAVPOTFOVAHLHAPITGSGSKTKVPAAYAAQGYKVLVLNPSVAATLGF 240  
 241 AYMSKAHGDIPNIRGTGRTITGAPITSTYGVKFLADGGCGGAYDIIICDECHSTDSTT 300  
 241 AYMSKAHGDIPNIRGTGRTITGAPITSTYGVKFLADGGCGGAYDIIICDECHSTDSTT 300  
 301 ILGIGTVLDOAETAGARLVLATATPGSVTPHPNIEEVALSTGRIPIFYKAIPIETI 360  
 301 ILGIGTVLDOAETAGARLVLATATPGSVTPHPNIEEVALSTGRIPIFYKAIPIETI 360  
 361 KGGHLLIFCHSKKCKDELAALSLGLNNAVAYRGLDVSVIPTSVDVIVVATDALMTGFT 420  
 361 KGGHLLIFCHSKKCKDELAALSLGLNNAVAYRGLDVSVIPTSVDVIVVATDALMTGFT 420  
 421 GDFSVIDCNTCTVTQVDFSLDPTFTTITVPODAYRSORRGRTGGRMGIIYRFVTPG 480  
 421 GDFSVIDCNTCTVTQVDFSLDPTFTTITVPODAYRSORRGRTGGRMGIIYRFVTPG 480  
 481 ERPSGMPDSSVLCBCYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCODHLEFWSVFTGL 540  
 481 ERPSGMPDSSVLCBCYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCODHLEFWSVFTGL 540  
 541 THIDAHFLSOTKQAGDNPPYLVAQVATVCARAQAPPPSWDQWKCLIRLKPRTLHGPTLL 600  
 541 THIDAHFLSOTKQAGDNPPYLVAQVATVCARAQAPPPSWDQWKCLIRLKPRTLHGPTLL 600  
 601 YRLGAVQNEVTTTHPIKYIMACHSADLEVTSWLVGGVLAALAAAYCLTGSVVIYGR 660  
 601 YRLGAVQNEVTTTHPIKYIMACHSADLEVTSWLVGGVLAALAAAYCLTGSVVIYGR 660  
 661 IILSKPAIIPDREVLVREFDEMBCASHLPYIEQGMOLAEQFKOAIQLLQATKQAEA 720  
 661 IILSKPAIIPDREVLVREFDEMBCASHLPYIEQGMOLAEQFKOAIQLLQATKQAEA 720  
 721 AAPVVESKRWTLFAFWAKHMMNFIISGQYLAGSLTLPGNPAIASLMAFTASITSLTQH 780  
 721 AAPVVESKRWTLFAFWAKHMMNFIISGQYLAGSLTLPGNPAIASLMAFTASITSLTQH 780  
 781 TLLFNILGWAAQALAPSASAFVAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 840  
 781 TLLFNILGWAAQALAPSASAFVAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 840  
 841 FKWMSGEMPESTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVAVQWNNLLAFASR 900  
 841 FKWMSGEMPESTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVAVQWNNLLAFASR 900  
 901 GNHVSPTHVVPESDAAARVTQILSLTITQLLKLHWINEDCSTPCSGSWLRDVMWDMIC 960  
 901 GNHVSPTHVVPESDAAARVTQILSLTITQLLKLHWINEDCSTPCSGSWLRDVMWDMIC 960  
 961 TVLTDFKTLWQSKLLPRLCPVPPFCORGYKGVWRGDMOTTCPCGAQITGHVKNXGMR 1020  
 961 TVLTDFKTLWQSKLLPRLCPVPPFCORGYKGVWRGDMOTTCPCGAQITGHVKNXGMR 1020  
 1021 IVGPRTCSTNWTGTPINAYTTGCTPPSPAPNYSRALMEVAAEYVEVTRVGDHFVVTGM 1080  
 1021 IVGPRTCSTNWTGTPINAYTTGCTPPSPAPNYSRALMEVAAEYVEVTRVGDHFVVTGM 1080  
 1081 TTDNVKPCQVPAPEFFTEVDGVRHLHRYAPACKPLLRBEVTFIVGUNQYLVGSQLPCEPE 1140  
 1081 TTDNVKPCQVPAPEFFTEVDGVRHLHRYAPACKPLLRBEVTFIVGUNQYLVGSQLPCEPE 1140  
 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASQLSAPSKATCTTHDSDPDAD 1200  
 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASQLSAPSKATCTTHDSDPDAD 1200  
 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPFRAM 1260  
 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPFRAM 1260

1261 PIWARDYNPPILLESKMDPDYVPPVHVCPLPPAKAPPIPPRRKRTVWLSSEVSALA 1320  
 1261 PIWARDYNPPILLESKMDPDYVPPVHVCPLPPAKAPPIPPRRKRTVWLSSEVSALA 1320  
 1321 ELATKTFGSSSESAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEPEGDPDLSDGSW 1380  
 1321 ELATKTFGSSSESAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEPEGDPDLSDGSW 1380  
 1381 STVSEASESDVCCMSYTWGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRAS 1440  
 1381 STVSEASESDVCCMSYTWGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRAS 1440  
 1441 LRQKKTVDRLQVLDHRYDVLKEMKASTVAKLLSVEEACKLTPPHSARSFGYCAK 1500  
 1441 LRQKKTVDRLQVLDHRYDVLKEMKASTVAKLLSVEEACKLTPPHSARSFGYCAK 1500  
 1501 DVRNLSSKANVHRSVWKDLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 1501 DVRNLSSKANVHRSVWKDLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 1561 VRVCEKMALVDVSTLPOAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620  
 1561 VRVCEKMALVDVSTLPOAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620  
 1621 TVTENDIRVBESIYQCCDLAPEARQALRSLSLTERLYIGGPLTNSKGQNGYRRCRASGVL 1680  
 1621 TVTENDIRVBESIYQCCDLAPEARQALRSLSLTERLYIGGPLTNSKGQNGYRRCRASGVL 1680  
 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740  
 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740  
 1741 SAPGDPKPKBYDELITSCSSNSVVAHDASGKVYLTTRDPTTPLARAAMETARHTPVN 1800  
 1741 SAPGDPKPKBYDELITSCSSNSVVAHDASGKVYLTTRDPTTPLARAAMETARHTPVN 1800  
 1801 SWLGNIIWYAPTLWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIQR 1860  
 1801 SWLGNIIWYAPTLWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIQR 1860  
 1861 HGLSAFSLHSPGEINRVASCLRLKGVPLRVHRHARSVRARLLSOGGRAATCGKYL 1920  
 1861 HGLSAFSLHSPGEINRVASCLRLKGVPLRVHRHARSVRARLLSOGGRAATCGKYL 1920  
 1921 NNAVTKLKLPIPAASQDLSSWVAGYSGGDIYHSLSRARPRWFMCLLLLSVGVI 1980  
 1921 NNAVTKLKLPIPAASQDLSSWVAGYSGGDIYHSLSRARPRWFMCLLLLSVGVI 1980  
 1981 LLPNR 1985  
 1981 LLPNR 1985

RESULT 3  
 US-09-539-601-6  
 ; Sequence 6, Application US/09539601C  
 ; Patent No. 6630343  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bartenschlager, Ralf FW  
 ; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
 ; FILE REFERENCE: all sequences  
 ; CURRENT APPLICATION NUMBER: US/09/539,601C  
 ; CURRENT FILING DATE: 2001-08-30  
 ; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
 ; EARLIER FILING DATE: 1999-04-03  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 2201  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 US-09-539-601-6



Query Match		100.0%; Score 10462; DB 4; Length 2201;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 1984; Conservative 1; Mismatches		0; Indels 0; Gaps 0;	
QY	1	MAPTAYSOQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSF	276
DB	217	LAPITAYSOQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSF	276
QY	61	GSKTLAGKGPITOMYTNVDDQVLGWAQAPPCARSITPCTCGSSDL	120
DB	277	GSKTLAGKGPITOMYTNVDDQVLGWAQAPPCARSITPCTCGSSDL	336
QY	121	GDSRGSLLSPRPVSYLKGSSGCLLCPGSHAVGIFRAAVCTRGVAKA	180
DB	337	GDSRGSLLSPRPVSYLKGSSGCLLCPGSHAVGIFRAAVCTRGVAKA	396
QY	181	RSVPFTDQAEATAGARLVLATATPPGSSVTPHPNIEEVALSSTG	240
DB	397	RSVPFTDQAEATAGARLVLATATPPGSSVTPHPNIEEVALSSTG	456
QY	241	AYMSKAHGIDPNIRGTITGAPITVSTYKFLADGCGGGAYDIIIC	300
DB	457	AYMSKAHGIDPNIRGTITGAPITVSTYKFLADGCGGGAYDIIIC	516
QY	301	ILGIGTVLDQAEATAGARLVLATATPPGSSVTPHPNIEEVALSSTG	360
DB	517	ILGIGTVLDQAEATAGARLVLATATPPGSSVTPHPNIEEVALSSTG	576
QY	361	KGRHLLFCCHSKKCDLAKLSGLGNAYAYRGLDVSIPTSGDVIV	420
DB	577	KGRHLLFCCHSKKCDLAKLSGLGNAYAYRGLDVSIPTSGDVIV	636
QY	421	GDSDSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVRSORRGR	480
DB	637	GDSDSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVRSORRGR	696
QY	481	ERPSGMPDSSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLPVC	540
DB	697	ERPSGMPDSSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLPVC	756
QY	541	THIDAHFLSKQAGNFPYLVAYQATVCARAQAPPSPDQMKCLIRL	600
DB	757	THIDAHFLSKQAGNFPYLVAYQATVCARAQAPPSPDQMKCLIRL	816
QY	601	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTTSTWLVGVLAA	660
DB	817	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTTSTWLVGVLAA	876
QY	661	IILSGKPAIIPDREVLRYRFBDEMEECASHLPYIEQGVLAEPKQ	720
DB	877	IILSGKPAIIPDREVLRYRFBDEMEECASHLPYIEQGVLAEPKQ	936
QY	721	AAPVVEKWRTELEAFWAKHWNFIISGIVLAGLSTLPGNPATIAS	780
DB	937	AAPVVEKWRTELEAFWAKHWNFIISGIVLAGLSTLPGNPATIAS	996
QY	781	TLLFNILGCVAAQLAPPSSAFAVGAGIAGAVGSGIGLVLDILAG	840
DB	997	TLLFNILGCVAAQLAPPSSAFAVGAGIAGAVGSGIGLVLDILAG	1056
QY	841	EKVMSEMPSTEDVNLPAIISPGALVGVVCAAILRRHVGPGEVAV	900
DB	1057	EKVMSEMPSTEDVNLPAIISPGALVGVVCAAILRRHVGPGEVAV	1116
QY	901	GNHVSPTHVPSDAAARVTQILSSLTITQLKRLHWINEDCSTPCS	960
DB	1117	GNHVSPTHVPSDAAARVTQILSSLTITQLKRLHWINEDCSTPCS	1176
QY	961	TVLTFDKTWLQSKLLPRLPGVFPFSCQRCYKGVWRGDGIMQTTCP	1020
DB	1177	TVLTFDKTWLQSKLLPRLPGVFPFSCQRCYKGVWRGDGIMQTTCP	1236
QY	1021	IVGPRCTSNWHTGCTPINAYTTGCTPSPAPNYSRALMRVAAEE	1080

RESULT 4

US-09-539-601-15

; Sequence 15, Application US/09539601C

; Patent No. 6630343

; GENERAL INFORMATION:

DB	1237	IVGPRCTSNWHTGCTPINAYTTGCTPSPAPNYSRALMRVAAEE	1296
QY	1081	TTDNVKCPQCPAPAEFFTEVDGRLHYAPACKLLREEVTLVGLNO	1140
DB	1297	TTDNVKCPQCPAPAEFFTEVDGRLHYAPACKLLREEVTLVGLNO	1356
QY	1141	PDVAVLTSMLTDPDPSHITAEATKRLARGSPSLASSASQSLKAT	1200
DB	1357	PDVAVLTSMLTDPDPSHITAEATKRLARGSPSLASSASQSLKAT	1416
QY	1201	LIEANLLWRQBMGNITRVSEKENVILDSPEPLQAEDEREVSVAE	1260
DB	1417	LIEANLLWRQBMGNITRVSEKENVILDSPEPLQAEDEREVSVAE	1476
QY	1261	PIWARPDPNPLLESWKDPDYPVPPVHGCPLPPAKAPPIPPRRKR	1320
DB	1477	PIWARPDPNPLLESWKDPDYPVPPVHGCPLPPAKAPPIPPRRKR	1536
QY	1321	ELATKTFGSSSSAVDSGTATASPDQSDGSDGSDVESYSMPPLGE	1380
DB	1537	ELATKTFGSSSSAVDSGTATASPDQSDGSDGSDVESYSMPPLGE	1596
QY	1381	STVSEEAESDVCCMSYTWGALITPCAEEETKLINALSNSLLRHNL	1440
DB	1597	STVSEEAESDVCCMSYTWGALITPCAEEETKLINALSNSLLRHNL	1656
QY	1441	LROKKTFTDLQVLDHRYDLVKEMKAKASTVAKALLSVEBEACKLT	1500
DB	1657	LROKKTFTDLQVLDHRYDLVKEMKAKASTVAKALLSVEBEACKLT	1716
QY	1501	DVRNLASKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEK	1560
DB	1717	DVRNLASKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEK	1776
QY	1561	VRYCEKMALVDVVTSLPOAVMGSSYFGVSPGQVFLVNAWKAKCP	1620
DB	1777	VRYCEKMALVDVVTSLPOAVMGSSYFGVSPGQVFLVNAWKAKCP	1836
QY	1621	TVTENDIRVEESIYQCCDLAPEARQAIIRSLTERLYIGGPLTNSK	1680
DB	1837	TVTENDIRVEESIYQCCDLAPEARQAIIRSLTERLYIGGPLTNSK	1896
QY	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICRSAGTQ	1740
DB	1897	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICRSAGTQ	1956
QY	1741	SAPPGDPKPEVDLELITSCSSNVSAHDASGKRVYLTTRDPTTPL	1800
DB	1957	SAPPGDPKPEVDLELITSCSSNVSAHDASGKRVYLTTRDPTTPL	2016
QY	1801	SWLGNIIIMVAPTILWARMILMTHFFSILLAQEQLEKALDCQIYG	1860
DB	2017	SWLGNIIIMVAPTILWARMILMTHFFSILLAQEQLEKALDCQIYG	2076
QY	1861	HGLSAPLSHSYSGPEINRVASCLRLGVPLRWHRARSVRARLLS	1920
DB	2077	HGLSAPLSHSYSGPEINRVASCLRLGVPLRWHRARSVRARLLS	2136
QY	1921	NWAVRTKLLTPIPAASQDLSSWFWAGVSGGDIYHLSIRARPRFW	1980
DB	2137	NWAVRTKLLTPIPAASQDLSSWFWAGVSGGDIYHLSIRARPRFW	2196
QY	1981	LLPNR 1985	
DB	2197	LLPNR 2201	



```
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2201
; TYPE: PR7
; ORGANISM: Hepatitis C virus
; US-09-539-601-15

Query Match      100.0%; Score 10462; DB 4; Length 2201;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPITAYSQTRGLGCGIITSLTGRDRNQVEGEVQVWSTATQSFATCNGVCWTVYHGA 60
Db :
217 LAPITAYSQTRGLGCGIITSLTGRDRNQVEGEVQVWSTATQSFATCNGVCWTVYHGA 276

Qy 61 GSKTLAGEKGPITOMYTNVDODLVGQWQAPPGARSLTPTCTCGSSDLVLTREHADVIVRRR 120
Db :
277 GSKTLAGEKGPITOMYTNVDODLVGQWQAPPGARSLTPTCTCGSSDLVLTREHADVIVRRR 336

Qy 121 GDSRGSLLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVESMETTM 180
Db :
337 GDSRGSLLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVESMETTM 396

Qy 181 RSPVFTDNSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAGQKVLVLPNSVAATLGF 240
Db :
397 RSPVFTDNSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAGQKVLVLPNSVAATLGF 456

Qy 241 AYMSKAGHIDNIRGTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTSTT 300
Db :
457 AYMSKAGHIDNIRGTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTSTT 516

Qy 301 ILGIGTVLDQAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 360
Db :
517 ILGIGTVLDQAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 576

Qy 361 KGRHLLIFCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420
Db :
577 KGRHLLIFCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 636

Qy 421 GDFDSVIDCNTVQTQVDFSLDPTFTTETTVQDAYSRSORRGRTGGRMGIIYRFVTPG 480
Db :
637 GDFDSVIDCNTVQTQVDFSLDPTFTTETTVQDAYSRSORRGRTGGRMGIIYRFVTPG 696

Qy 481 ERPSGMFDSVLCBICYDAGCAWYELTPAETSVRRLAYLNTPLGLPVCODHLFWESVFTGL 540
Db :
697 ERPSGMFDSVLCBICYDAGCAWYELTPAETSVRRLAYLNTPLGLPVCODHLFWESVFTGL 756

Qy 541 THIDAHFLSQTKQAGDPPYLVAVQATVCARAQAPPPSDQMWKCLIRLKPTELHGPTPL 600
Db :
757 THIDAHFLSQTKQAGDPPYLVAVQATVCARAQAPPPSDQMWKCLIRLKPTELHGPTPL 816

Qy 601 YRLGAVQNEVTTTHPIKTYINWACMSADLEVTVSTWVLVGGVLAALAAVCLTGTGVIIVGR 660
Db :
817 YRLGAVQNEVTTTHPIKTYINWACMSADLEVTVSTWVLVGGVLAALAAVCLTGTGVIIVGR 876

Qy 661 IILSGKPAIIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKATIGLLQTATKQAEA 720
Db :
877 IILSGKPAIIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKATIGLLQTATKQAEA 936

Qy 721 AAPVVESKRWLTLEAFWAKHMNFISGIIYLAGLSTLPGNPAIASMAFTASITSPLTQ 780
Db :
937 AAPVVESKRWLTLEAFWAKHMNFISGIIYLAGLSTLPGNPAIASMAFTASITSPLTQ 996

Qy 781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGAGALVA 840
Db :
997 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGAGALVA 1056

841 FKWSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAGVQWNRLLIAFASR 900
1057 FKWSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAGVQWNRLLIAFASR 1116

Qy 901 GNHVSPTHYPESDAAARVTQILSSLLITQLKELHQMINEDCSTPCSGSLRDVWDWIC 960
Db :
1117 GNHVSPTHYPESDAAARVTQILSSLLITQLKELHQMINEDCSTPCSGSLRDVWDWIC 1176

Qy 961 TVLTDFTKMLQSKLLPRLPGVPPFSCORGKYGWVRGDIQCTTCPCGAQITGHVKNVSGMR 1020
Db :
1177 TVLTDFTKMLQSKLLPRLPGVPPFSCORGKYGWVRGDIQCTTCPCGAQITGHVKNVSGMR 1236

Qy 1021 IGVPRCTSNTHGTFPFINAYTTGCTPSPAPNYSRALWRVAEEYEVTRVGDFFHYVTGM 1080
Db :
1237 IGVPRCTSNTHGTFPFINAYTTGCTPSPAPNYSRALWRVAEEYEVTRVGDFFHYVTGM 1296

Qy 1081 TTDNVKPCQVPAPPEFTEVDGVLHRYAPACKLLREEVTFVLGLNQYLVGSLQPCPE 1140
Db :
1297 TTDNVKPCQVPAPPEFTEVDGVLHRYAPACKLLREEVTFVLGLNQYLVGSLQPCPE 1356

Qy 1141 PDVAVLTSMLTDPSHITAEAKRLARGSPSPSLASSASQSLAPSLKATCTTTRHDSPPAD 1200
Db :
1357 PDVAVLTSMLTDPSHITAEAKRLARGSPSPSLASSASQSLAPSLKATCTTTRHDSPPAD 1416

Qy 1201 LIEANLWRQEMGNNITRVESENKVVILDSFEPLOAEDEDEREVSPABILRSRKFPPAM 1260
Db :
1417 LIEANLWRQEMGNNITRVESENKVVILDSFEPLOAEDEDEREVSPABILRSRKFPPAM 1476

Qy 1261 PIWARPDPYNPPLLESWKDPDYPVPVHGCPLPRAKAPPIPPRRKRTVILSESTVSSALA 1320
Db :
1477 PIWARPDPYNPPLLESWKDPDYPVPVHGCPLPRAKAPPIPPRRKRTVILSESTVSSALA 1536

Qy 1321 ELATKTFGSSSADVSTGATASDPQSDDGSDGSDVESYSSMPPEGECPDPLSDGWS 1380
Db :
1537 ELATKTFGSSSADVSTGATASDPQSDDGSDGSDVESYSSMPPEGECPDPLSDGWS 1596

Qy 1381 STVSEEAESDVVCCSMSTYWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 1440
Db :
1597 STVSEEAESDVVCCSMSTYWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 1656

Qy 1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEACKLTPPHSARSFEGYAK 1500
Db :
1657 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEACKLTPPHSARSFEGYAK 1716

Qy 1501 DVRLSKKAVNHNRSVMKDLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLIVFPDLG 1560
Db :
1717 DVRLSKKAVNHNRSVMKDLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLIVFPDLG 1776

Qy 1561 VRVCEKMALYDVWSTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 1620
Db :
1777 VRVCEKMALYDVWSTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 1836

Qy 1621 TVTENDIRVBEESIYQCCDLAPPEARQATRSUTERLYIGPLTNSKQNGVRCRCASGVLIT 1680
Db :
1837 TVTENDIRVBEESIYQCCDLAPPEARQATRSUTERLYIGPLTNSKQNGVRCRCASGVLIT 1896

Qy 1681 TSCNTLITCYLKAACAAKALQDCTMLVCGDDLVIWICESAGTOEDEASLRAFTTEAMTRY 1740
Db :
1897 TSCNTLITCYLKAACAAKALQDCTMLVCGDDLVIWICESAGTOEDEASLRAFTTEAMTRY 1956

Qy 1741 SAPPGDPPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLARAAMETARHTFVN 1800
Db :
1957 SAPPGDPPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLARAAMETARHTFVN 2016

Qy 1801 SWLGNIIYAPTLLWARMILMTHFPFIIILAQOLEKALDCQIYGACYSIEPLDLQIQL 1860
Db :
2017 SWLGNIIYAPTLLWARMILMTHFPFIIILAQOLEKALDCQIYGACYSIEPLDLQIQL 2076

Qy 1861 HGLSAFSLHSYSGPEINRVSCLKLVPPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920
Db :
2077 HGLSAFSLHSYSGPEINRVSCLKLVPPPLRVWRHARSVRARLLSQGGRAATCGKYL 2136
```

QY 1921 NWAURTKLKTPIPAASQLDSSWVAGYSGGDIYHLSRARPWFMMCLLLSVGVGIY 1380  
DB 2137 NWAURTKLKTPIPAASQLDSSWVAGYSGGDIYHLSRARPWFMMCLLLSVGVGIY 2196  
QY 1981 LLENR 1985  
DB 2197 LLENR 2201  
RESULT 5  
US-09-539-601-3  
; Sequence 3, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartschlagel, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 3010  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-3  
Query Match 100.0%; Score 10462; DB 4; Length 3010;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIIITSLTGRDRNQVEGEVQVVSSTATQSFLLATCVNGCWTVYHGA 60  
DB 1026 LAPITAYSQOTRGLLCIIITSLTGRDRNQVEGEVQVVSSTATQSFLLATCVNGCWTVYHGA 1085  
QY 61 GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPGARSLTPCTCGSSDLVIVTRHADVIPVRR 120  
DB 1086 GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPGARSLTPCTCGSSDLVIVTRHADVIPVRR 1145  
QY 121 GDSRGLLSPRPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
DB 1146 GDSRGLLSPRPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 1205  
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
DB 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 1265  
QY 241 AYMSKAHGIDPNIRTCVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 300  
DB 1266 AYMSKAHGIDPNIRTCVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 1325  
QY 301 ILGIGTVLQDAETAGARLVVLAATATPGSVTVPHPNIEEVALSSTGEIPEYKAIPIETI 360  
DB 1326 ILGIGTVLQDAETAGARLVVLAATATPGSVTVPHPNIEEVALSSTGEIPEYKAIPIETI 1385  
QY 361 KGRGHLIFCHSKKKCKDELAALSGLGINAVAYRGLDVSVIPTSGDVIVVATDALTMTGT 420  
DB 1386 KGRGHLIFCHSKKKCKDELAALSGLGINAVAYRGLDVSVIPTSGDVIVVATDALTMTGT 1445  
QY 421 GFDSDVIDCNTCTQTVDFSLDPTFTIETTTPQDAVRSORRGRTGRGMGIYRFVTPG 480  
DB 1446 GFDSDVIDCNTCTQTVDFSLDPTFTIETTTPQDAVRSORRGRTGRGMGIYRFVTPG 1505  
QY 481 ERPSGMFDSVLCCEYDAGCAWYELTAPETSVRRLAYLNTPLPVCQDHLHFWESVFTGL 540  
DB 1506 ERPSGMFDSVLCCEYDAGCAWYELTAPETSVRRLAYLNTPLPVCQDHLHFWESVFTGL 1565  
QY 541 THIDAFLSQTQAGDNEPVLVAYQATVCARAQAPPPSDMKLIRLKPTLHGPTPLL 600  
DB 1566 THIDAFLSQTQAGDNEPVLVAYQATVCARAQAPPPSDMKLIRLKPTLHGPTPLL 1625

QY 601 YRLGAVQNEVTHPTTKYIMACMSADLEVVVTWLVGGVLAALAAAYCLTTSVVLVGR 660  
DB 1626 YRLGAVQNEVTHPTTKYIMACMSADLEVVVTWLVGGVLAALAAAYCLTTSVVLVGR 1685  
QY 661 IILSGKPAIIPDREVLYRPFDEMEECASHLPYIEQGMQLAEQFKQKAIGLLOATKQAEA 720  
DB 1686 IILSGKPAIIPDREVLYRPFDEMEECASHLPYIEQGMQLAEQFKQKAIGLLOATKQAEA 1745  
QY 721 AAPVSESKWRTLEAFWAKHWNFIISGIQYIAGLSTLPGNPAIASLMAFTASITSLTQTH 780  
DB 1746 AAPVSESKWRTLEAFWAKHWNFIISGIQYIAGLSTLPGNPAIASLMAFTASITSLTQTH 1805  
QY 781 TLLFNILGGWVAQAAPPAAAFVAGAGTAGAVSGIGLKVLDIILGAGVAGVALVA 840  
DB 1806 TLLFNILGGWVAQAAPPAAAFVAGAGTAGAVSGIGLKVLDIILGAGVAGVALVA 1865  
QY 841 FKVMGEMFSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLLIAFASR 900  
DB 1866 FKVMGEMFSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLLIAFASR 1925  
QY 901 GNVSPHVVPSDAAARVTOILSSLTITQLLKRHLQWINEDECSTPCSGSWLRDWDWIC 960  
DB 1926 GNVSPHVVPSDAAARVTOILSSLTITQLLKRHLQWINEDECSTPCSGSWLRDWDWIC 1985  
QY 961 TVLTDFKTWLOSLLPRLPGVPPFCORGYKVMRGDMQTTCPGGAQITGHVKGSMR 1020  
DB 1986 TVLTDFKTWLOSLLPRLPGVPPFCORGYKVMRGDMQTTCPGGAQITGHVKGSMR 2045  
QY 1021 IVPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 1080  
DB 2046 IVPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPILREEVTVLGLNOYLVGSQLPCEPE 1140  
DB 2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPILREEVTVLGLNOYLVGSQLPCEPE 2165  
QY 1141 PDVAVLTSMLTDSHITAEAKRLARGPPSPASSASQLSAPSLKATCTTRHDSPDAD 1200  
DB 2166 PDVAVLTSMLTDSHITAEAKRLARGPPSPASSASQLSAPSLKATCTTRHDSPDAD 2225  
QY 1201 LIEANLLMRQEMGNITRVESKNVILDSFEPLQAEDEEREVSFPAELRRSRKPPRAM 1260  
DB 2226 LIEANLLMRQEMGNITRVESKNVILDSFEPLQAEDEEREVSFPAELRRSRKPPRAM 2285  
QY 1261 PIWARDYNNPPLLESKWDYVPPVHGCPLPPAKAPPITPPRRKKTIVLSESTVSALA 1320  
DB 2286 PIWARDYNNPPLLESKWDYVPPVHGCPLPPAKAPPITPPRRKKTIVLSESTVSALA 2345  
QY 1321 ELATKTFGSESSAVDSGTATASPDQSDDDAGSDVESYSSMPPLEGEPEGDPLSDGSW 1380  
DB 2346 ELATKTFGSESSAVDSGTATASPDQSDDDAGSDVESYSSMPPLEGEPEGDPLSDGSW 2405  
QY 1381 STVSEASBDDVCCMSYTWGALITPCAABETKLPINALNSLLRHNLLVYATTSRAS 1440  
DB 2406 STVSEASBDDVCCMSYTWGALITPCAABETKLPINALNSLLRHNLLVYATTSRAS 2465  
QY 1441 LRQKKVTFDLQVLDHVDVLEKMAKASTVAKALLSVEEACKLTPPHSARSKFYGA 1500  
DB 2466 LRQKKVTFDLQVLDHVDVLEKMAKASTVAKALLSVEEACKLTPPHSARSKFYGA 2525  
QY 1501 DVNLSKKAVNHRSVWKDLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
DB 2526 DVNLSKKAVNHRSVWKDLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585  
QY 1561 VRVCEHMAIYDVVSTLPOAVMGSSYFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 1620  
DB 2586 VRVCEHMAIYDVVSTLPOAVMGSSYFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS 2645  
QY 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKQNGCVRRCRASGVL 1680  
DB 2646 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKQNGCVRRCRASGVL 2705

QY 1681 TSCGNLTTCYLKAAACAAKQDCTMLVCGDDLVVWICESAGTOEDASIRAFTEAMTRY 1740  
Db 2706 TSCGNLTTCYLKAAACAAKQDCTMLVCGDDLVVWICESAGTOEDASIRAFTEAMTRY 2765  
QY 1741 SAPGDPKPKPYDLELITSCSSNSVVAHDASGKRVYVLTTRDPTTPLARAAWETARHTPVN 1800  
Db 2766 SAPGDPKPKPYDLELITSCSSNSVVAHDASGKRVYVLTTRDPTTPLARAAWETARHTPVN 2825  
QY 1801 SWLGNIIWAPTILWARMILMTHFFSILLIAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
Db 2826 SWLGNIIWAPTILWARMILMTHFFSILLIAQBLEKALDCQIYGACYSIEPLDLQIIQRL 2885  
QY 1861 HGLSAFSLHSYSGEINRVASCLKGLVPLRVWRHRSVRARLLSOGGAAATCGKYL 1920  
Db 2886 HGLSAFSLHSYSGEINRVASCLKGLVPLRVWRHRSVRARLLSOGGAAATCGKYL 2945  
QY 1921 NWAVRTKLTPIPAASOLDLSSWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGII 1980  
Db 2946 NWAVRTKLTPIPAASOLDLSSWFVAGYSGGDIYHLSRARPRFWMCLLLLSVGVGII 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

## RESULT 6

US-10-029-907-3  
; Sequence 3, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 882  
; OTHER INFORMATION: Xaa is Lys or Arg  
; NAME/KEY: VARIANT  
; LOCATION: 1489  
; OTHER INFORMATION: Xaa is Leu

US-10-029-907-3

Query Match 99.8%; Score 10442; DB 4; Length 2201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPITAYSQQRGLGCIITSLTGRDRNQVEGEVQVWSTATQSFATCVCVGVVYHGA 60  
Db 217 LAPITAYSQQRGLGCIITSLTGRDRNQVEGEVQVWSTATQSFATCVCVGVVYHGA 276  
QY 61 GSKTLAGKGPITQMYTNVDQDLVGWQAPPGARSITPCTCGSSDLVYTRHADVIPVRR 120  
Db 277 GSKTLAGKGPITQMYTNVDQDLVGWQAPPGARSITPCTCGSSDLVYTRHADVIPVRR 336  
QY 121 GDSRGSLLSPRPVSYLKSSGGPILLCPCHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 337 GDSRGSLLSPRPVSYLKSSGGPILLCPCHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 396  
QY 181 RSPVFTDNSSPPAVPQTQFVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVNPVSAATLGF 240  
Db 397 RSPVFTDNSSPPAVPQTQFVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVNPVSAATLGF 456

QY 241 AYMSKAHGIIDPNIITGVRTITTTGAPITTYSTYKFLADGGCSGGAYDIIICDECHSTDSTT 300  
Db 457 AYMSKAHGIIDPNIITGVRTITTTGAPITTYSTYKFLADGGCSGGAYDIIICDECHSTDSTT 516  
QY 301 ILGIGTVLDOAETAGARLVLATATPPGSVTVPHNPTEEVALSSTGEIPFPYKAIPIETI 360  
Db 517 ILGIGTVLDOAETAGARLVLATATPPGSVTVPHNPTEEVALSSTGEIPFPYKAIPIETI 576  
QY 361 KGRSHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIPTSGDVIIVATDALMTGFT 420  
Db 577 KGRSHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIPTSGDVIIVATDALMTGFT 636  
QY 421 GDFPSVDCNCTVQTQVDFSLDPTFTTITTVPODAVRSORRGTRGRMGIIIRFVTPG 480  
Db 637 GDFPSVDCNCTVQTQVDFSLDPTFTTITTVPODAVRSORRGTRGRMGIIIRFVTPG 696  
QY 481 ERPSGMFDSVLCBCYDAGCAWVELTPAETSVRIRAYLNTPLGPVQCQDHLFWESVFTGL 540  
Db 697 ERPSGMFDSVLCBCYDAGCAWVELTPAETSVRIRAYLNTPLGPVQCQDHLFWESVFTGL 756  
QY 541 THIDAHFLSQTQAGDNFPYLVAYQATVCARAAQAPPSPSWDMKCLIRLKPTELHGPTLL 600  
Db 757 THIDAHFLSQTQAGDNFPYLVAYQATVCARAAQAPPSPSWDMKCLIRLKPTELHGPTLL 816  
QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTSGVIVGR 660  
Db 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTSGVIVGR 876  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKOKAIGLLQATKQAEA 720  
Db 877 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKOKAIGLLQATKQAEA 936  
QY 721 AAPVYESKWRTELEAFWAKHMMNFIISQIYLAGSLTLPENPAIASIMAFASITPLTQ 780  
Db 937 AAPVYESKWRTELEAFWAKHMMNFIISQIYLAGSLTLPENPAIASIMAFASITPLTQ 996  
QY 781 TLLFNILGGWVAQALAPPSAASAEVAGIAGAAVGSIGLGVLDIILAGYAGAGALVA 840  
Db 997 TLLFNILGGWVAQALAPPSAASAEVAGIAGAAVGSIGLGVLDIILAGYAGAGALVA 1056  
QY 841 FKVMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILAFASR 900  
Db 1057 FKVMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILAFASR 1116  
QY 901 GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDECSPTCSGSLRDVDMWIC 960  
Db 1117 GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDECSPTCSGSLRDVDMWIC 1176  
QY 961 TVLTDKFTWLSKLLPRLPGVPPFSCORGKGVWRGDMQTTCCGNOITGHVKNCSMR 1020  
Db 1177 TVLTDKFTWLSKLLPRLPGVPPFSCORGKGVWRGDMQTTCCGNOITGHVKNCSMR 1236  
QY 1021 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAEEVEVTRVGDHYVTGM 1080  
Db 1237 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAEEVEVTRVGDHYVTGM 1296  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRLHRYAPACKPLLEEVTVLVLNQYLVGSLPCEPE 1140  
Db 1297 TTDNVKCPQVPAPEFFTEVDGVRLHRYAPACKPLLEEVTVLVLNQYLVGSLPCEPE 1356  
QY 1141 PDVAVLTSMLTDPSHITAETAKRLARGSPSLASSASQLSAPSLKATCTTTRHDSPAD 1200  
Db 1357 PDVAVLTSMLTDPSHITAETAKRLARGSPSLASSASQLSAPSLKATCTTTRHDSPAD 1416  
QY 1201 LIEANLLWRQEMGMGNITRVESENKVWILDSEFPLQAEDEEREVSVPAILRRSRKFFPRAM 1260  
Db 1417 LIEANLLWRQEMGMGNITRVESENKVWILDSEFPLQAEDEEREVSVPAILRRSRKFFPRAM 1476  
QY 1261 PIWARPYNPLLESWKDPDYPVPVHVHGCPLPPAKAPPIPPPRKRRTVWLSESTVSSALA 1320  
Db 1477 PIWARPYNPLLESWKDPDYPVPVHVHGCPLPPAKAPPIPPPRKRRTVWLSESTVSSALA 1536  
QY 1321 ELATKTFGSSSESSAVDSGTATASPDQDDGAGSDVESYSSMPPLEGEPCDPLSDGSW 1380

Db 1537 ELATKTFGSSSESAVDSGTATSPDQSDSDAGSDVESYSSPPLEGEGFGDPLSDGSW 1596  
QY 1381 STVSEASESDVCCSMSTYWTGALITPCAAABETKLPINALNSLLRHNNLVYATTSSAS 1440  
Db 1597 STVSEASESDVCCSMSTYWTGALITPCAAABETKLPINALNSLLRHNNLVYATTSSAS 1656  
QY 1441 LROKKTFTDFRLOVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
Db 1657 LROKKTFTDFRLOVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK 1716  
QY 1501 DVNLSSKAVNHRSVWKDLEDETETPDTITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
Db 1717 DVNLSSKAVNHRSVWKDLEDETETPDTITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1776  
QY 1561 VRCEKMALYDVVSTLTPQAVMGSSYGHQSPGQRFVFLVNAWAKKCPMGFAVDTTRCFDS 1620  
Db 1777 VRCEKMALYDVVSTLTPQAVMGSSYGHQSPGQRFVFLVNAWAKKCPMGFAVDTTRCFDS 1836  
QY 1621 TVTENDIRVEESIQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGVRRCRASGVLT 1680  
Db 1837 TVTENDIRVEESIQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGVRRCRASGVLT 1896  
QY 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEDEASLRAFTTEAMTRY 1740  
Db 1897 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEDEASLRAFTTEAMTRY 1956  
QY 1741 SAPPGPPKPEYDELELITSCSNVSVNAHDASGRVYVLTDPDTPPLARAAWETARHTPVN 1800  
Db 1957 SAPPGPPKPEYDELELITSCSNVSVNAHDASGRVYVLTDPDTPPLARAAWETARHTPVN 2016  
QY 1801 SWLGNIIIMVAPTLLWARMILMTHFFSILLAOEQLKALDCQIYACYSIEPLDLPOIIOQL 1860  
Db 2017 SWLGNIIIMVAPTLLWARMILMTHFFSILLAOEQLKALDCQIYACYSIEPLDLPOIIOQL 2076  
QY 1861 HGLSAPSLHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGRAATCGKLYF 1920  
Db 2077 HGLSAPSLHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGRAATCGKLYF 2136  
QY 1921 NNAVTRKLTPTIPAAASQLDLSWFAVAGYSGGDIYHLSLRPRPFWMCCLLLSVGVGIY 1980  
Db 2137 NNAVTRKLTPTIPAAASQLDLSWFAVAGYSGGDIYHLSLRPRPFWMCCLLLSVGVGIY 2196  
QY 1981 LLENR 1985  
Db 2197 LLENR 2201

RESULT 7  
US-09-539-601-24  
; Sequence 24, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartschlagel, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-24

Query Match 99.7%; Score 10435; DB 4; Length 1985;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1979; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIITSLTGRDRNVEGEVQVSTATQSFATCNGVCWTVYHGA 60

Db 1 MAPITAYSQOTRGLLCIITSLTGRDRNVEGEVQVSTATQSFATCNGVCWTVYHGA 60  
QY 61 GSKTLGAPKPIPTOMTNTVDQDLVGMQAPPGARSLLPTCTCGSSDLYLVTBHADVIVPVR 120  
Db 61 GSKTLGAPKPIPTOMTNTVDQDLVGMQAPPGARSLLPTCTCGSSDLYLVTBHADVIVPVR 120  
QY 121 GDSRGLILSPRPVSYLKGSSGGPLLCPGSHAVGIFRAAIVCTRCGVAKAVDFVPVSEMETTM 180  
Db 121 GDSRGLILSPRPVSYLKGSSGGPLLCPGSHAVGIFRAAIVCTRCGVAKAVDFVPVSEMETTM 180  
QY 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQYKVLVLPNSVAATLFGF 240  
Db 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQYKVLVLPNSVAATLFGF 240  
QY 241 AYNKKAHGDIPNIRTVGRTITTCAPITYSTYGFADGGCGGGAYDIIICDECHSTDSTT 300  
Db 241 AYNKKAHGDIPNIRTVGRTITTCAPITYSTYGFADGGCGGGAYDIIICDECHSTDSTT 300  
QY 301 ILGIGTVLQDAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPEYKKAIPLETI 360  
Db 301 ILGIGTVLQDAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPEYKKAIPLETI 360  
QY 361 KGRHILIFCHSKKKCCDELAALKSLGGLNVAAYYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
Db 361 KGRHILIFCHSKKKCCDELAALKSLGGLNVAAYYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
QY 421 GPDFSDVIDCNTVOTVDFSLDPTFTIETTVQDQAVSQRGRGTHGMYRFPVTPG 480  
Db 421 GPDFSDVIDCNTVOTVDFSLDPTFTIETTVQDQAVSQRGRGTHGMYRFPVTPG 480  
QY 481 ERPSGMFDSVLCCEYDAGCAWYELTPAETSVRRLAYLNTPLGVPQDHLFEWESVFTGL 540  
Db 481 ERPSGMFDSVLCCEYDAGCAWYELTPAETSVRRLAYLNTPLGVPQDHLFEWESVFTGL 540  
QY 541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWDMQWKCLIRLKPTLHGPTPL 600  
Db 541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWDMQWKCLIRLKPTLHGPTPL 600  
QY 601 YRLGAVQNEVTHPTTKYIMACMSADLEWVTSWLVGVLAALAAAYCLTTCSSVIVGR 660  
Db 601 YRLGAVQNEVTHPTTKYIMACMSADLEWVTSWLVGVLAALAAAYCLTTCSSVIVGR 660  
QY 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEQKQKAIQLQATKQAEA 720  
Db 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEQKQKAIQLQATKQAEA 720  
QY 721 AAPVVEKWRTEAFWAKHMMNFIQIYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db 721 AAPVVEKWRTEAFWAKHMMNFIQIYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
QY 781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA 840  
Db 781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA 840  
QY 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVSGEGAVQMMNRLIAFASR 900  
Db 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVSGEGAVQMMNRLIAFASR 900  
QY 901 GNVHSPTHVPSDAAAARVTQILSSLTITQLKRLHQWINECDSTPCSSGSLRDVWDWIC 960  
Db 901 GNVHSPTHVPSDAAAARVTQILSSLTITQLKRLHQWINECDSTPCSSGSLRDVWDWIC 960  
QY 961 TVLTDFKWLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCGAOITGHVKNGSMR 1020  
Db 961 TVLTDFKWLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCGAOITGHVKNGSMR 1020  
QY 1021 IVGPRTCSNTHGTFPINAITYTTGCTPSPAPNYSRALWRVAABEYVEVTRGVGFHYVTGM 1080  
Db 1021 IVGPRTCSNTHGTFPINAITYTTGCTPSPAPNYSRALWRVAABEYVEVTRGVGFHYVTGM 1080  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPILREEVTFVLGNQILVGSQLPCEPE 1140

Db 1081 TTDDVKPCQCPAPEBFTTEVDGVLHRYAPACKPLLEEVTFLVGLNQLVGSQLPCEPE 1140  
QY 1141 PVAVLITMLTDPSSHITATKRLARGSPPLSSASSASQLSAPSUKATCTTRHDSPPAD 1200  
Db 1141 PVAVLITMLTDPSSHITATKRLARGSPPLSSASSASQLSAPSUKATCTTRHDSPPAD 1200  
QY 1201 LIEANLLMQEMGNGNITRVESENKVVILDSFPLQAEEDEREVSVAEILRRSRKPPRAM 1260  
Db 1201 LIEANLLMQEMGNGNITRVESENKVVILDSFPLQAEEDEREVSVAEILRRSRKPPRAM 1260  
QY 1261 PIWAPDYNNPPLLESWKDPPYVPPVHGCPLPPAKAPPIPPRRKRTVVLSSESTVSSALA 1320  
Db 1261 PIWAPDYNNPPLLESWKDPPYVPPVHGCPLPPAKAPPIPPRRKRTVVLSSESTVSSALA 1320  
QY 1321 ELATITFGSSSESAVDSGTATASPOPSDDGAGSDVESYSSMPLEGGPDLSDGSW 1380  
Db 1321 ELATITFGSSSESAVDSGTATASPOPSDDGAGSDVESYSSMPLEGGPDLSDGSW 1380  
QY 1381 STVSEASEDVVCCSMTYWTGALITPCAABETKLPINALSNLLRHHNLVATTSSRAS 1440  
Db 1381 STVSEASEDVVCCSMTYWTGALITPCAABETKLPINALSNLLRHHNLVATTSSRAS 1440  
QY 1441 LRQKVTDRLOVLDHDDHVDLVKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
Db 1441 LRQKVTDRLOVLDHDDHVDLVKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
QY 1501 DVRLNSSKAVNHRSVWKDLLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIYFPDLG 1560  
Db 1501 DVRLNSSKAVNHRSVWKDLLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIYFPDLG 1560  
QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQVSPQORVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
Db 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQVSPQORVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
QY 1621 TVTENDIRVERSIYOCCLDAPPEARQAIRSLTERLYIGGLTNSKQNCGYRRCRASGVLT 1680  
Db 1621 TVTENDIRVERSIYOCCLDAPPEARQAIRSLTERLYIGGLTNSKQNCGYRRCRASGVLT 1680  
QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLAFTEAMTRY 1740  
Db 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLAFTEAMTRY 1740  
QY 1741 SAPPGDPKPYDLELITSCSSNVSVADHASKRVYLYLTRPTTPPARAAWETARHTPVN 1800  
Db 1741 SAPPGDPKPYDLELITSCSSNVSVADHASKRVYLYLTRPTTPPARAAWETARHTPVN 1800  
QY 1801 SWLGNIIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
Db 1801 SWLGNIIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
QY 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYL 1920  
Db 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYL 1920  
QY 1921 NWAVRTKLUTPIPAASQDLSSWFVAGYSGGDIYHSLSRARPPWFVWMLLLLSVGVGIY 1980  
Db 1921 NWAVRTKLUTPIPAASQDLSSWFVAGYSGGDIYHSLSRARPPWFVWMLLLLSVGVGIY 1980  
QY 1981 LLPNR 1985  
Db 1981 LLPNR 1985

RESULT 8  
US-09-539-601-27  
; Sequence 27, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30

; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 3010  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-27

Query Match 99.7%; Score 10432; DB 4; Length 3010;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1978; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLGCIIITSLTGRDRNOVEGEVQVSTATOSFLATCVNGVCWTYYHGA 60  
Db 1026 LAPITAYSQOTRGLLGCIIITSLTGRDRNOVEGEVQVSTATOSFLATCVNGVCWTYYHGA 1085  
QY 61 GSKTLAGPKGPITOMYTNVDQDLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 120  
Db 1086 GSKTLAGPKGPITOMYTNVDQDLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 1145  
QY 121 GDSRGSLLSPRPVSVYKSSGGPLLCPSGHANGVIFRAAVCTRGVAKAVDFVPVESMETTM 180  
Db 1146 GDSRGSLLSPRPVSVYKSSGGPLLCPSGHANGVIFRAAVCTRGVAKAVDFVPVESMETTM 1205  
QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQYKVLVNPVSAATLGGF 240  
Db 1206 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQYKVLVNPVSAATLGGF 1265  
QY 241 AYMSKAHGIDNIRTVGRTITITGAPITVYSTYKFLADGGCGSGAYDIIICDECHSTDSTT 300  
Db 1266 AYMSKAHGIDNIRTVGRTITITGAPITVYSTYKFLADGGCGSGAYDIIICDECHSTDSTT 1325  
QY 301 ILGIGTVLDQAEAGARLVLATATPPGCVTVPHNIEEVALSSTGEIPFYGKAIPETI 360  
Db 1326 ILGIGTVLDQAEAGARLVLATATPPGCVTVPHNIEEVALSSTGEIPFYGKAIPETI 1385  
QY 361 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVVATDALMTGFT 420  
Db 1386 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVVATDALMTGFT 1445  
QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVRSQRGRGTGRMGIRFYVTPG 480  
Db 1446 GDFDSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVRSQRGRGTGRMGIRFYVTPG 1505  
QY 481 ERPSGMFDSVLCBCYDAGCAWYELTFAETSVRIRAYLINTPLGVPQCQHLBFWSVFTGL 540  
Db 1506 ERPSGMFDSVLCBCYDAGCAWYELTFAETSVRIRAYLINTPLGVPQCQHLBFWSVFTGL 1565  
QY 541 THIDAHFLSQTQKAGDNFPYIVAYQATVCARAQAPPSPSWDMWKCLIRLKPFLHGTPTLL 600  
Db 1566 THIDAHFLSQTQKAGDNFPYIVAYQATVCARAQAPPSPSWDMWKCLIRLKPFLHGTPTLL 1625  
QY 601 YRLGAVONEVTTTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSAWIVGR 660  
Db 1626 YRLGAVONEVTTTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSAWIVGR 1685  
QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIIGLQATKQAEA 720  
Db 1686 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIIGLQATKQAEA 1745  
QY 721 AAPVWESKRTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSLPTTQH 780  
Db 1746 AAPVWESKRTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSLPTTQH 1805  
QY 781 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAGALVA 1865  
QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLIIFASR 900  
Db 1866 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLIIFASR 1925

QY	901	GNVSPETHVPSDAAARVTQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960
Db	1926	GNHVSPTHVPSDAAARVTQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	1985
QY	961	TVLTDEKTMLOSKLLRLPGVPFSCORGKGVWVRGDMQTCPCGAQITGHVKNWSMR	1020
Db	1986	TVLTDFKTMLOSKLLRLPGVPFSCORGKGVWVRGDMQTCPCGAQITGHVKNWSMR	2045
QY	1021	IVGPRCSNTHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDVHYVTGM	1080
Db	2046	IVGPRCSNTHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDVHYVTGM	2105
QY	1081	TTDNVCKPCQVPAPEFTEVDGVRHLRYAPACKPLLEBEVTVFLVGLNOYLVSQLECEPE	1140
Db	2106	TTDDVRCPCQVPAPEFTEVDGVRHLRYAPACKPLLEBEVTVFLVGLNOYLVSQLECEPE	2165
QY	1141	PDVAVLTSMLTDPSSHITAETAKERLARGSPPLASSASOLSAPLSKATCTTTRHDSFDAD	1200
Db	2166	PDVAVLTSMLTDPSSHITAETAKERLARGSPPLASSASOLSAPLSKATCTTTRHDSFDAD	2225
QY	1201	LIEANLWRQEMGNITRVSENVKVLDSFEPLQAEEDEREVSVPALILRRSRKPPRAM	1260
Db	2226	LIEANLWRQEMGNITRVSENVKVLDSFEPLQAEEDEREVSVPALILRRSRKPPRAM	2285
QY	1261	PIWARDPNPLLESWMKDDPYVPPVHGCPLPAKAPPPIPPRRKRTVVLSESTVSALA	1320
Db	2286	PIWARDPNPLLESWMKDDPYVPPVHGCPLPAKAPPPIPPRRKRTVVLSESTVSALA	2345
QY	1321	ELATKTFGSSSAVDSGTATASPDQSDGSDGSDVESYSSMPPLGEPEGDPDLSGWS	1380
Db	2346	ELATKTFGSSSAVDSGTATASPDQSDGSDGSDVESYSSMPPLGEPEGDPDLSGWS	2405
QY	1381	STVSEASESDVCCSMSTYTGTALITPCAAEETKLPINALSNSLLRHNLVYATTSRSAS	1440
Db	2406	STVSEASESDVCCSMSTYTGTALITPCAAEETKLPINALSNSLLRHNLVYATTSRSAS	2465
QY	1441	LROKKTFTDLQVLDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSFEGYAK	1500
Db	2466	LROKKTFTDLQVLDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSFEGYAK	2525
QY	1501	DVRNLSSKANVNHISVMKDLLEDTEPTDITIMAKNEVFVCQPEKGRKPARLIVFPDLG	1560
Db	2526	DVRNLSSKANVNHISVMKDLLEDTEPTDITIMAKNEVFVCQPEKGRKPARLIVFPDLG	2585
QY	1561	VRVCEKVALXDVVSTLPQAVMGSSYFGYSPQRVEFLVNAWAKKCPMGFAYDTRCFDS	1620
Db	2586	VRVCEKVALXDVVSTLPQAVMGSSYFGYSPQRVEFLVNAWAKKCPMGFAYDTRCFDS	2645
QY	1621	TVTENDIRVESIYQCCDLAPARQAIRSLTERLYIGPLTNSKGNGYRRCRASGLT	1680
Db	2646	TVTENDIRVESIYQCCDLAPARQAIRSLTERLYIGPLTNSKGNGYRRCRASGLT	2705
QY	1681	TSCGNLTLCYLKAAACRAAKLQCTMLVCGDDLWVLCESAGTQDEASLRAFTTEAMTRY	1740
Db	2706	TSCGNLTLCYLKAAACRAAKLQCTMLVCGDDLWVLCESAGTQDEASLRAFTTEAMTRY	2765
QY	1741	SAPPGDPPKPEYDLELITSCSNVSVAHDSAGKRVYILTRDPTPLARAAMETARHTPVN	1800
Db	2766	SAPPGDPPKPEYDLELITSCSNVSVAHDSAGKRVYILTRDPTPLARAAMETARHTPVN	2825
QY	1801	SWLGNINMYATLWARMILMTHFSLIAQOLEKALDCQYAGCYSIEPLDLQIOLR	1860
Db	2826	SWLGNINMYATLWARMILMTHFSLIAQOLEKALDCQYAGCYSIEPLDLQIOLR	2885
QY	1861	HGLSAFSLHSYSPGEINRVASCLRLKGLVPPPLRVWRHARSVRARLLSQGGRAAATCGKYL	1920
Db	2886	HGLSAFSLHSYSPGEINRVASCLRLKGLVPPPLRVWRHARSVRARLLSQGGRAAATCGKYL	2945
QY	1921	NWAVRTKULPTIPAASQLDLSWVAGYSGDDIYHSLSRAPRPFMCLLLLSVGVGII	1980
Db	2946	NWAVRTKULPTIPAASQLDLSWVAGYSGDDIYHSLSRAPRPFMCLLLLSVGVGII	3005
QY	1981	LLPNR	1985
Db	3006	LLPNR	3010
RESULT 9			
US-09-539-601-18			
; Sequence 18, Application US/09539601C			
; Patent No. 6630343			
; GENERAL INFORMATION:			
; APPLICANT: Bartenschlager, Ralf FW			
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System			
; FILE REFERENCE: all sequences			
; CURRENT APPLICATION NUMBER: US/09/539,601C			
; CURRENT FILING DATE: 2001-08-30			
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY			
; EARLIER FILING DATE: 1999-04-03			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: Patent in Ver. 2.1			
; SEQ ID NO 18			
; LENGTH: 1985			
; TYPE: PRT			
; ORGANISM: Hepatitis C virus			
US-09-539-601-18			
Query Match 99.6%; Score 10425; DB 4; Length 1985;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1977; Conservative 4; Mismatches 4; Indels 0; Gaps 0;			
QY	1	MAPITAYSQQTGRLGCIITSLGRDRNQVEGEVQVVSATQSFATCVNGVCWTVYHGA	60
Db	1	MAPITAYSQQTGRLGCIITSLGRDRNQVEGEVQVVSATQSFATCVNGVCWTVYHGA	60
QY	61	GSKTLAGPKPIQMTYNVDOLVGCQAPPGASLTCTCGSSDLXLYLTHADVIVRRR	120
Db	61	GSKTLAGPKPIQMTYNVDOLVGCQAPPGASLTCTCGSSDLXLYLTHADVIVRRR	120
QY	121	GDGRGSLSPRPVSYLYKGGSGGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180
Db	121	GDGRGSLSPRPVSYLYKGGSGGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180
QY	181	RSVPFTDNNSPFPVQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKVILNPSVAATLFGF	240
Db	181	RSVPFTDNNSPFPVQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKVILNPSVAATLFGF	240
QY	241	AYMSKAHGIQDNRTGVTITTCGAPITYSTYKFLADGCGGGAYDIIICDECHSTDSST	300
Db	241	AYMSKAHGIQDNRTGVTITTCGAPITYSTYKFLADGCGGGAYDIIICDECHSTDSST	300
QY	301	ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNIEEVALSSTGPIPFYKKAIPETI	360
Db	301	ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNIEEVALSSTGPIPFYKKAIPETI	360
QY	361	KGGSHLIFCHSKKFCDELAIAKLSGLGNNAVAYRGLDVSVIPTSQGDVIVATDALMTGFT	420
Db	361	KGGSHLIFCHSKKFCDELAIAKLSGLGNNAVAYRGLDVSVIPTSQGDVIVATDALMTGFT	420
QY	421	GDPSVVDICNTCVTQTVDFSLDPTFTTETTTVPQDAVRSORRGTGRGMGIYRFVTPG	480
Db	421	GDPSVVDICNTCVTQTVDFSLDPTFTTETTTVPQDAVRSORRGTGRGMGIYRFVTPG	480
QY	481	ERPSGMFDSVSLCECYDAGCAWYELTPAETSRLRAYLNTPLGFPVCDHLEFWESVFTGL	540
Db	481	ERPSGMFDSVSLCECYDAGCAWYELTPAETSRLRAYLNTPLGFPVCDHLEFWESVFTGL	540
QY	541	THIDAHFLSQTKQAGDNFPYLVAYQATVCARQAQPPPSDQMKCLIRLKPTELHGTPLL	600
Db	541	THIDAHFLSQTKQAGDNFPYLVAYQATVCARQAQPPPSDQMKCLIRLKPTELHGTPLL	600
QY	601	YRLGAVQNEVTTTHPIITKYIMACMSADLEVTSTVWLVGGVLAALAAAYCLTTGVSIVVGR	660
Db	601	YRLGAVQNEVTTTHPIITKYIMACMSADLEVTSTVWLVGGVLAALAAAYCLTTGVSIVVGR	660

661 IILSGKPAIPDREVLYRFEDEMEECASHLPYIEQGMQLAEQKQKAIQLQATATKQAEA 720  
Db  
661 IILSGKPAIPDREVLYRFEDEMEECASHLPYIEQGMQLAEQKQKAIQLQATATKQAEA 720  
QY  
721 AAPVBSKWRITLBAFAWKHMWNFISGQYIAGLSTLPGNPAIASLMAFTASITSPITTOH 780  
Db  
721 AAPVBSKWRITLBAFAWKHMWNFISGQYIAGLSTLPGNPAIASLMAFTASITSPITTOH 780  
QY  
781 TILLENILGGMWAQALAPPSAASAFVAGAGTAGAAGVIGLGVKLVLDILAGYGAGVAGALVA 840  
Db  
781 TILLENILGGMWAQALAPPSAASAFVAGAGTAGAAGVIGLGVKLVLDILAGYGAGVAGALVA 840  
QY  
841 FKVMGEMPSDETLVNLPAILSPGALVVGWVCAALRRHVGPGEAGVOMNRLIAFASR 900  
Db  
841 FKVMGEMPSDETLVNLPAILSPGALVVGWVCAALRRHVGPGEAGVOMNRLIAFASR 900  
QY  
901 GNVHSPTHVPSDDAAARVTQILSSITITQLLKRHLQWINEDECSPTCSGSLRDLVMDWIC 960  
Db  
901 GNVHSPTHVPSDDAAARVTQILSSITITQLLKRHLQWINEDECSPTCSGSLRDLVMDWIC 960  
QY  
961 TVLTDPKWLQSKLLPRLPGLVPPFFSCORGYKGVWRGDGIMQITTCPCGAOITGHVKNXGMR 1020  
Db  
961 TVLTDPKWLQSKLLPRLPGLVPPFFSCORGYKGVWRGDGIMQITTCPCGAOITGHVKNXGMR 1020  
QY  
1021 IVGPRTCSTWHTGTFPINAATYTGCTPPSPAPNYSRALWVAEEVVEVTRVGDHRYVTGM 1080  
Db  
1021 IVGPRTCSTWHTGTFPINAATYTGCTPPSPAPNYSRALWVAEEVVEVTRVGDHRYVTGM 1080  
QY  
1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLLEEVTVLVLNOLYVLSQPCPBE 1140  
Db  
1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLLEEVTVLVLNOLYVLSQPCPBE 1140  
QY  
1141 PDVAVLTSMLTDPSSHITAEKRLARGSPPLSSASSASOLSAPSLKATCTTRHDSPAD 1200  
Db  
1141 PDVAVLTSMLTDPSSHITAEKRLARGSPPLSSASSASOLSAPSLKATCTTRHDSPAD 1200  
QY  
1201 LIEANLLWRQMGNGNITRVESENKVVILDSPEPLQAEDEREVSPAEILRRSRKFFPRAM 1260  
Db  
1201 LIEANLLWRQMGNGNITRVESENKVVILDSPEPLQAEDEREVSPAEILRRSRKFFPRAM 1260  
QY  
1261 PIWARPDPNPLLSKWDKDPDVPVPHGCPPLPPAKAPPPIPPRRKRTVVLSESTVSSALA 1320  
Db  
1261 PIWARPDPNPLLSKWDKDPDVPVPHGCPPLPPAKAPPPIPPRRKRTVVLSESTVSSALA 1320  
QY  
1321 ELATKTFGSSSESSAVDSGTATASDPQSDGSDGSDVESYMSMPLEGEQDPLSDGSW 1380  
Db  
1321 ELATKTFGSSSESSAVDSGTATASDPQSDGSDGSDVESYMSMPLEGEQDPLSDGSW 1380  
QY  
1381 STVSEASEDVPCCSMTWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
Db  
1381 STVSEASEDVPCCSMTWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
QY  
1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACKLTPPHSARSFYGAK 1500  
Db  
1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACKLTPPHSARSFYGAK 1500  
QY  
1501 DVRNLSSKANVHIBSVKMLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLLVFPDPLG 1560  
Db  
1501 DVRNLSSKANVHIBSVKMLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLLVFPDPLG 1560  
QY  
1561 VRVCEKMALXDVNSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620  
Db  
1561 VRVCEKMALXDVNSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620  
QY  
1621 TVTENDIRVEESIVQCCDLAPAPQAIRSLTERLYIGPLTNSKGQNGYRRCRASGVLT 1680  
Db  
1621 TVTENDIRVEESIVQCCDLAPAPQAIRSLTERLYIGPLTNSKGQNGYRRCRASGVLT 1680  
QY  
1681 TSCGNTLTCVYLKAAACRAAKLODCTMLVCGDDLLVVICESAGTOEDEASIRAFTEAMTRY 1740  
Db  
1681 TSCGNTLTCVYLKAAACRAAKLODCTMLVCGDDLLVVICESAGTOEDEASIRAFTEAMTRY 1740  
QY  
1741 SAPPGDPPKPEYDLELITSCSSNVSVAHDSAGKRVYVLTDRPTTTLARAAWEIARHTPVN 1800

Db 1741 SAPPGDPPKPEYDLELITSCSSNVSVAHDSAGKRVYVLTDRPTTTLARAAWEIARHTPVN 1800  
QY 1801 SWLGNLIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCOIQYACYSIEPLDLPQIQR 1860  
Db 1801 SWLGNLIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCOIQYACYSIEPLDLPQIQR 1860  
QY 1861 HGLSAFSLHSYSPGEINRVASCLRLKLGVPPLRVHRHARSVRARLLSQGGRATATCKYLF 1920  
Db 1861 HGLSAFSLHSYSPGEINRVASCLRLKLGVPPLRVHRHARSVRARLLSQGGRATATCKYLF 1920  
QY 1921 NNAVTRKLTLPPIAASQDLSSWFWAGVSGGDIYHSLSRARPRMFWCMLLLSVGVGIY 1980  
Db 1921 NNAVTRKLTLPPIAASQDLSSWFWAGVSGGDIYHSLSRARPRMFWCMLLLSVGVGIY 1980  
QY 1981 LILPNR 1985  
Db 1981 LILPNR 1985  
RESULT 10  
US-09-539-601-21  
; Sequence 21, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 3010  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-21  
Query Match 99.6%; Score 10422; DB 4; Length 3010;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1976; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MAPITAYSQQTRGLLGLCIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCTVYHGA 60  
Db 1026 LAPITAYSQQTRGLLGLCIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCTVYHGA 1085  
QY 61 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVRHADVIPVRRR 120  
Db 1086 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVRHADVIPVRRR 1145  
QY 121 GDSRGSLLSRPVSYLKSGSGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
Db 1146 GDSRGSLLSRPVSYLKSGSGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 1205  
QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQGVKVLVLPNSVAATLGRG 240  
Db 1206 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQGVKVLVLPNSVAATLGRG 1265  
QY 241 AYMSKAHIGDPINRITGVRTITTCAPITYSTYKFLADGGCGSGGAYDIIICDECHSTDSTT 300  
Db 1266 AYMSKAHIGDPINRITGVRTITTCAPITYSTYKFLADGGCGSGGAYDIIICDECHSTDSTT 1325  
QY 301 ILGTGTVLDOAETAGARLVVLATATPPGCVTVPHPNIEEVALSTGPIFYGKAIPETI 360  
Db 1326 ILGTGTVLDOAETAGARLVVLATATPPGCVTVPHPNIEEVALSTGPIFYGKAIPETI 1385  
QY 361 KGRHILIFCHSKKKKDELAALKSLGLNAVYVYRGLDVSVIPTSGDVIVVATATLMTGFT 420  
Db 1386 KGRHILIFCHSKKKKDELAALKSLGLNAVYVYRGLDVSVIPTSGDVIVVATATLMTGFT 1445  
QY 421 GDFDSVIDNCVTQTVDFSLDPTFTTITTTVPQDAVSRQRGRGTGRGMGIYRFVTPG 480



Db 1446 GDFSDVDCNTCVTQVDFSLDFTFFIETTTVPQDAVSRSQRGRGRGMYIRVFTG 1505  
QY 481 ERPSGMFSDSVLCECYDAGCAWYELTPAETSVELRAYLNTPLPVQDHLFEWESVFTGL 540  
Db 1506 ERPSGMFSDSVLCECYDAGCAWYELTPAETSVELRAYLNTPLPVQDHLFEWESVFTGL 1565  
QY 541 THIDAFLSOTKQAGNFPYLVAYQATVCARAQAPPSWDQWKCLIRLKPITLHGPTLL 600  
Db 1566 THIDAFLSOTKQAGNFPYLVAYQATVCARAQAPPSWDQWKCLIRLKPITLHGPTLL 1625  
QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTGTVLWVGVAALAAAYCLTTGSSVIVGR 660  
Db 1626 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTGTVLWVGVAALAAAYCLTTGSSVIVGR 1685  
QY 661 IILSGKPAIIPDREVLRYEPDEMEBCASHLPYIEQGMQALAEQFKQKAIQLQTATQAAEA 720  
Db 1686 IILSGKPAIIPDREVLRYEPDEMEBCASHLPYIEQGMQALAEQFKQKAIQLQTATQAAEA 1745  
QY 721 AAPVSEKWTLEAFWAKHWNFTSGYOYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db 1746 AAPVSEKWTLEAFWAKHWNFTSGYOYLAGLSTLPGNPAIASLMAFTASITSPLTQH 1805  
QY 781 TLLFNILGWWAQAAPPASAFAVAGIAGAAVGSIGLGVLDIILAGYAGAGALVA 840  
Db 1806 TLLFNILGWWAQAAPPASAFAVAGIAGAAVGSIGLGVLDIILAGYAGAGALVA 1865  
QY 841 FKVMGMPSTEDLVNLLPAILSPGALVGVVCAAILRHVGPGEAGVQWNNRLIAPASR 900  
Db 1866 FKVMGMPSTEDLVNLLPAILSPGALVGVVCAAILRHVGPGEAGVQWNNRLIAPASR 1925  
QY 901 GNEVSPHYYPESDAAARVQILSSLTITOLLKELHOWINEDCSTPCSGWLRDWDWIC 960  
Db 1926 GNEVSPHYYPESDAAARVQILSSLTITOLLKELHOWINEDCSTPCSGWLRDWDWIC 1985  
QY 961 TVLTDFTKTLQSKLLPRLPGVPFSCQYKGVWRGDMQTTCCPGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFTKTLQSKLLPRLPGVPFSCQYKGVWRGDMQTTCCPGAQITGHVKNQSMR 2045  
QY 1021 IGVPRTCSTNWHGTFPFINAYTTGCTSPAPNYGRALWRVAABEYVEVTRVGDHYHTGM 1080  
Db 2046 IGVPRTCSTNWHGTFPFINAYTTGCTSPAPNYGRALWRVAABEYVEVTRVGDHYHTGM 2105  
QY 1081 TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPLLEEVTVLGLNQYLWGSOLPCPE 1140  
Db 2106 TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPLLEEVTVLGLNQYLWGSOLPCPE 2165  
QY 1141 PDVAULTSMULTDPSHITAETAKRRLARGSPPLASSASQSLAPSLKATCTTRHDSPAD 1200  
Db 2166 PDVAULTSMULTDPSHITAETAKRRLARGSPPLASSASQSLAPSLKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWRQMGNGNITRVESENKVILDSPEPIQAEDEDEREVSPABILRRSRKFFRAM 1260  
Db 2226 LIEANLLWRQMGNGNITRVESENKVILDSPEPIQAEDEDEREVSPABILRRSRKFFRAM 2285  
QY 1261 PIWARPDPNPLLESWDPDVPVPHGCPPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 2286 PIWARPDPNPLLESWDPDVPVPHGCPPLPPAKAPPIPPRRKRTVVLSESTVSSALA 2345  
QY 1321 ELATKTFGSSSSAVDSGTATASPDQSDGSDGSDVESMSPPLEGEPDPLSDGSW 1380  
Db 2346 ELATKTFGSSSSAVDSGTATASPDQSDGSDGSDVESMSPPLEGEPDPLSDGSW 2405  
QY 1381 STVSEASEDVVCCSMSTWTGALITPCAABETKLPINALSNSLIRHNLVYATTSRAS 1440  
Db 2406 STVSEASEDVVCCSMSTWTGALITPCAABETKLPINALSNSLIRHNLVYATTSRAS 2465  
QY 1441 LRQKVTDFRLQVLDHVRDLVKEMKASTVKAKLLSVEACKLTPPHSARSKFYGAK 1500  
Db 2466 LRQKVTDFRLQVLDHVRDLVKEMKASTVKAKLLSVEACKLTPPHSARSKFYGAK 2525  
QY 1501 DVNRLSSKAVNHRSVWKDLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLG 1560

Db 2526 DVNRLSSKAVNHRSVWKDLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLG 2585  
QY 1561 VRVCEKWAYDYVSTLTPQAVMGSSYGFQYSPGORVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Db 2586 VRVCEKWAYDYVSTLTPQAVMGSSYGFQYSPGORVEFLVNAWKAKKCPMGFAYDTRCFDS 2645  
QY 1621 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRASGVL 1680  
Db 2646 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRASGVL 2705  
QY 1681 TSCGNLTLCYLKAAACRAKLODCTMLVCGDDLVI CESAGTOEDEASLRAFTEAMTRY 1740  
Db 2706 TSCGNLTLCYLKAAACRAKLODCTMLVCGDDLVI CESAGTOEDEASLRAFTEAMTRY 2765  
QY 1741 SAPPGPPKPEYDLELITSCSSNVSAHDASGRVYLLTRDPTTPLARAAWETARHTPVN 1800  
Db 2766 SAPPGPPKPEYDLELITSCSSNVSAHDASGRVYLLTRDPTTPLARAAWETARHTPVN 2825  
QY 1801 SWLGNTIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 1860  
Db 2826 SWLGNTIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 2885  
QY 1861 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYL 1920  
Db 2886 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYL 2945  
QY 1921 NWAVRTKLTLPITPAASQDLSWVFAGYSGGDIYHLSLRARPRFWMCLLLSVGVGIY 1980  
Db 2946 NWAVRTKLTLPITPAASQDLSWVFAGYSGGDIYHLSLRARPRFWMCLLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

## RESULT 11

US-09-539-601-30  
; Sequence 30, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartschlagel, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-30

Query Match 99.5%; Score 10412; DB 4; Length 1985;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1974; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAPITAYSQOQTRGLLGCIIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCMVYHGA 60  
Db 1 MAPITAYSQOQTRGLLGCIIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCMVYHGA 60  
QY 61 GSKTLAGKGPITOMYTNVDQDLVGWQAPPCARSITPCTCGSSDLVLTVRHADVIPVRRR 120  
Db 61 GSKTLAGKGPITOMYTNVDQDLVGWQAPPCARSITPCTCGSSDLVLTVRHADVIPVRRR 120  
QY 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVESMETTM 180  
Db 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVESMETTM 180  
QY 181 RSPVFTDNSSPPAPVQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVNPSSVAATLFG 240



181 RSPVFTDNGSPAPVQTFQVAHLHPTGSGKSTKVEAAVAGQYKVLVNLNPSVAATLGF 240  
241 AYMSKAHGDENIRITGVRITTCGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
241 AYMSKAHGDENIRIGVRITTCGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
301 ILGIGTVLQDAETAGARLVLTATATPPGVTVPHPNIEEVALSSGTGEIPFYKAIPIETI 360  
301 ILGIGTVLQDAETAGARLVLTATATPPGVTVPHPNIEEVALSSGTGEIPFYKAIPIETI 360  
361 KGRHLIFCHSKKCKDELAAGSLGCLNNAVYRGLDVSVIPTSGDVIIVVATDALMTGFT 420  
361 KGRHLIFCHSKKCKDELAAGSLGCLNNAVYRGLDVSVIPTSGDVIIVVATDALMTGFT 420  
421 GDFSDVIDCNTCTQTVDFSLDPTFTIETTTVPQAVSRQRGRGGRMGIYFVTPG 480  
421 GDFSDVIDCNTCTQTVDFSLDPTFTIETTTVPQAVSRQRGRGGRMGIYFVTPG 480  
481 ERPSGMFSSVLCYDAGCAWYELTPAETSVRLRAYLNTGTLVPCQDHLFEWESVFTGL 540  
481 ERPSGMFSSVLCYDAGCAWYELTPAETSVRLRAYLNTGTLVPCQDHLFEWESVFTGL 540  
541 THIDAHFISQTKQAGNPPYLVAQATVCARAQAPPPSWDQWKKLIRLKP TLHGFTPLL 600  
541 THIDAHFISQTKQAGNPPYLVAQATVCARAQAPPPSWDQWKKLIRLKP TLHGFTPLL 600  
601 YRLGAVQNEVTTTHPIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVVIYGR 660  
601 YRLGAVQNEVTTTHPIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVVIYGR 660  
661 IILSKPAIIPDREVLXREFDEMEBCASHLPYIEQGMQLAEQFKQKAILLQATKQABA 720  
661 IILSKPAIIPDREVLXREFDEMEBCASHLPYIEQGMQLAEQFKQKAILLQATKQABA 720  
721 AAPVVEKWTLEAPWAKHWNFTSGIYQIYLAGSLTPGNPAIASLMAFTASITSLTTOH 780  
721 AAPVVEKWTLEAPWAKHWNFTSGIYQIYLAGSLTPGNPAIASLMAFTASITSLTTOH 780  
781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 840  
781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 840  
841 FKVMSGEMPTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900  
841 FKVMSGELPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900  
901 GNHVSPTHYVPESDAAARVTOILSLTITQLKRLHWINEDCSTPCGSLRVDWDMIC 960  
901 GNHVSPTHYVPESDAAARVTOILSLTITQLKRLHWINEDCSTPCGSLRVDWDMIC 960  
961 TVLTDFTLWQSKLLPRLPGVPPFFSCORGKYGVRGDMOTTCPCGAQITGHVKNQSMR 1020  
961 TVLTDFTLWQSKLLPRLPGVPPFFSCORGKYGVRGDMOTTCPCGAQITGHVKNQSMR 1020  
1021 IVGPRTCSNTWHGTFPINAIVTGPCTSPAPNYGRALWRVAEEYVEVTRVGDPHYVTGM 1080  
1021 IVGPRTCSNTWHGTFPINAIVTGPCTSPAPNYGRALWRVAEEYVEVTRVGDPHYVTGM 1080  
1081 TTDNVKCPQVPAPEFTEVDGVRHLRYAPACKPLLEEVTVLVLNQYLVGSLQPCPE 1140  
1081 TTDNVKCPQVPAPEFTEVDGVRHLRYAPACKPLLEEVTVLVLNQYLVGSLQPCPE 1140  
1141 PDVAVLTSMLTDSHITAEAKRLARGSPSSASSASQSLAPSKATCTTRHSDPAD 1200  
1141 PDVAVLTSMLTDSHITAEAKRLARGSPSSASSASQSLAPSKATCTTRHSDPAD 1200  
1201 LIEANLLWRQMGNIITRVESENKVVILDSFEPLQAEDEDERVSVPAEILRRKRKPPRAM 1260  
1201 LIEANLLWRQMGNIITRVESENKVVILDSFEPLQAEDEDERVSVPAEILRRKRKPPRAM 1260  
1261 PIWARDYNPPLLESWKDPPYVPPVHGCPPLPAKAPPPIPPRRKRTVILSESTVSALA 1320  
1261 PIWARDYNPPLLESWKDPPYVPPVHGCPPLPAKAPPPIPPRRKRTVILSESTVSALA 1320

1321 ELATKTFGSSSANDSGTATATSPDQSDDGAGSDVESYSSMPLEGBEGDPLDGGSW 1380  
1321 ELATKTFGSSSANDSGTATATSPDQSDDGAGSDVESYSSMPLEGBEGDPLDGGSW 1380  
1381 STVSEASDGVCCSMSTYTWTGALITPCAABETKLPINALSNLLRHHNLVYATTSSAS 1440  
1381 STVSEASDGVCCSMSTYTWTGALITPCAABETKLPINALSNLLRHHNLVYATTSSAS 1440  
1441 LRQKVTFORLOVLDHVDLKEKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGA 1500  
1441 LRQKVTFORLOVLDHVDLKEKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGA 1500  
1501 DVRLNSSKAVNHRSVKDLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
1501 DVRLNSSKAVNHRSVKDLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKKCPMGFAYDTRCFDS 1620  
1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKKCPMGFAYDTRCFDS 1620  
1621 TVTENDIRVESIYOCCLDAPARQAIIRSLTERLYIGGLTNSKGNCGYRECRASGVLT 1680  
1621 TVTENDIRVESIYOCCLDAPARQAIIRSLTERLYIGGLTNSKGNCGYRECRASGVLT 1680  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
1741 SAPGDPKPKYDLELITSCSNVSVVAHDAGKRVYIYTRDPTTPPLARAAMTARHTPVN 1800  
1741 SAPGDPKPKYDLELITSCSNVSVVAHDAGKRVYIYTRDPTTPPLARAAMTARHTPVN 1800  
1801 SWLGNIIWYATPLWARMILMTHFFSILLAOBLEKALDCQIYGCYSTIEPLDLOIQL 1860  
1801 SWLGNIIWYATPLWARMILMTHFFSILLAOBLEKALDCQIYGCYSTIEPLDLOIQL 1860  
1861 HGLSAFSLHSPGEINRVASCLRLKLVPLVRVHRARSVRARLLSOGGRAATCGKYL 1920  
1861 HGLSAFSLHSPGEINRVASCLRLKLVPLVRVHRARSVRARLLSOGGRAATCGKYL 1920  
1921 NNAVTKLTPIPAASQDLSSWFVAGYSGDIIYHLSRARPWFVWMLLLSVGVGIY 1980  
1921 NNAVTKLTPIPAASQDLSSWFVAGYSGDIIYHLSRARPWFVWMLLLSVGVGIY 1980  
1981 LLPNR 1985  
1981 LLPNR 1985

## RESULT 12

US-09-539-601-33  
; Sequence 33, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 3010  
; TYPE: PRY  
; ORGANISM: Hepatitis C virus  
US-09-539-601-33

Query Match 99.5%; Score 10409; DB 4; Length 3010;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1973; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY	1	MAPITAYSQOTRGLLCIIITSITGRDRNVEGEVQVSTATOSFLATCNGVCWTYHGA	60	QY	1081	TTDNVCKPCQVPAPPEFFTEVDGVRHLRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	1140
Db	1026	LAPITAYSQOTRGLLCIIITSITGRDRNVEGEVQVSTATOSFLATCNGVCWTYHGA	1085	Db	2106	TTDNVCKPCQVPAPPEFFTEVDGVRHLRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	2165
QY	61	GSKTLAGPKPIOMYNTVNDODLVGQAPGARSITPCTCGSSDLVLRHADVIPVRR	120	QY	1141	PDVAVLTSMLTDPSSHITATKRRRLARGSPPSLASSASQLSAPSILKATCTTRHDSPDAD	1200
Db	1086	GSKTLAGPKPIOMYNTVNDODLVGQAPGARSITPCTCGSSDLVLRHADVIPVRR	1145	Db	2166	PDVAVLTSMLTDPSSHITATKRRRLARGSPPSLASSASQLSAPSILKATCTTRHDSPDAD	2225
QY	121	GDSRGLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFPVVPMETTM	180	QY	1201	LTEANLLWQEMGNGNTRVSEENKVVILDSFEPLOAEEDEREVSVAEILRRRKRKPPRAM	1260
Db	1146	GDSRGLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFPVVPMETTM	1205	Db	2226	LTEANLLWQEMGNGNTRVSEENKVVILDSFEPLOAEEDEREVSVAEILRRRKRKPPRAM	2285
QY	181	RSVFDTNDSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	240	QY	1261	PIWARDYNPPILLESWKDPDYVPPVVGCPPLPAKAPPIPPRRKRKTVLSESTVSALA	1320
Db	1206	RSVFDTNDSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	1265	Db	2286	PIWARDYNPPILLESWKDPDYVPPVVGCPPLPAKAPPIPPRRKRKTVLSESTVSALA	2345
QY	241	AYMSKAGHDIPNIRIGVRIITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST	300	QY	1321	ELATKTFGSESAVDSGTATASPDOPSDDGADGSDVESYSSMPPLEGEBPDGDLSDGSW	1380
Db	1266	AYMSKAGHDIPNIRIGVRIITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST	1325	Db	2346	ELATKTFGSESAVDSGTATASPDOPSDDGADGSDVESYSSMPPLEGEBPDGDLSDGSW	2405
QY	301	ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNIEEVALSGTGEIPFYKAIPIETI	360	QY	1381	STVSEASEDDVCCMSYTWGTALITPCAABETKLINALSNLLRHHNLVYATTGRSAS	1440
Db	1326	ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNIEEVALSGTGEIPFYKAIPIETI	1385	Db	2406	STVSEASEDDVCCMSYTWGTALITPCAABETKLINALSNLLRHHNLVYATTGRSAS	2465
QY	361	KGRHLIFCHSKKKKDELAALSGLGNLNAVAYRGLDVSVIPSGDVIVVATDALMTGFT	420	QY	1441	LROKKTFTORLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK	1500
Db	1386	KGRHLIFCHSKKKKDELAALSGLGNLNAVAYRGLDVSVIPSGDVIVVATDALMTGFT	1445	Db	2466	LROKKTFTORLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK	2525
QY	421	GDPSVLDNCTCVOTWDFSLDPTFTTTTPQDASVRSORRGRTGRGMGIYRFVTPG	480	QY	1501	DVRNLSKAVNHTRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560
Db	1446	GDPSVLDNCTCVOTWDFSLDPTFTTTTPQDASVRSORRGRTGRGMGIYRFVTPG	1505	Db	2526	DVRNLSKAVNHTRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	2585
QY	481	ERPSGMEFSSVCECYDAGCAWYBELTAPTSVRLRAYLNTPLPVCQDHLFEFVESVFTGL	540	QY	1561	VRYCEKMAVDVYSTLPQAVMGSSYGOYSPGQORVEPLVNAWAKAKCPMGFAYDTRCFDS	1620
Db	1506	ERPSGMEFSSVCECYDAGCAWYBELTAPTSVRLRAYLNTPLPVCQDHLFEFVESVFTGL	1565	Db	2586	VRYCEKMAVDVYSTLPQAVMGSSYGOYSPGQORVEPLVNAWAKAKCPMGFAYDTRCFDS	2645
QY	541	THIDAHF.SQTKQAGDNFPIYVAQVATCARAQAPPSWDQMWKCLRLKPTLHGPTLL	600	QY	1621	TVTENDIRVESIYQCCDLAPEARQAIIRSLTERLYIGGPLTNSKGQNCGRRCRASGVLT	1680
Db	1566	THIDAHF.SQTKQAGDNFPIYVAQVATCARAQAPPSWDQMWKCLRLKPTLHGPTLL	1625	Db	2646	TVTENDIRVESIYQCCDLAPEARQAIIRSLTERLYIGGPLTNSKGQNCGRRCRASGVLT	2705
QY	601	YRLGAVQNEVTTTHPIIKYINACMSADLEVVTSWLVGGVLAALAYCLTGSWIVGR	660	QY	1681	TSCGNTTUTCVLKAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTTEAMTRY	1740
Db	1626	YRLGAVQNEVTTTHPIIKYINACMSADLEVVTSWLVGGVLAALAYCLTGSWIVGR	1685	Db	2706	TSCGNTTUTCVLKAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTTEAMTRY	2765
QY	661	IILSGKPAIIPDREVLYREFDEMEBECASHLPYIEQGMQLABQFKOKAIGLLQTATKQAEA	720	QY	1741	SAPGDPKPKRYDIELITSCSSNVSVVAHDASGKRVYVLTTRDPTTPLARAAMETARHTPVN	1800
Db	1686	IILSGKPAIIPDREVLYREFDEMEBECASHLPYIEQGMQLABQFKOKAIGLLQTATKQAEA	1745	Db	2766	SAPGDPKPKRYDIELITSCSSNVSVVAHDASGKRVYVLTTRDPTTPLARAAMETARHTPVN	2825
QY	721	AAPVVEKWRTEAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH	780	QY	1801	SWLGNIIIMYAPTLWARMILMTHFTSILLAQBLEKALDCQIYGACYSIEPIJDLQIQR	1860
Db	1746	AAPVVEKWRTEAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH	1805	Db	2826	SWLGNIIIMYAPTLWARMILMTHFTSILLAQBLEKALDCQIYGACYSIEPIJDLQIQR	2885
QY	781	TLLFNILGGWAAQOLAPPSAASAFVAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA	840	QY	1861	HGLSAPSLHSYSGPEINRVASCLRKLGVPPLVWRHRRARSVRARLLSOGGAAATCGKYL	1920
Db	1806	TLLFNILGGWAAQOLAPPSAASAFVAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA	1865	Db	2886	HGLSAPSLHSYSGPEINRVASCLRKLGVPPLVWRHRRARSVRARLLSOGGAAATCGKYL	2945
QY	841	FKVMSGEMPTSEDLVNLPLAILSPGALVGVVCAAILRRHVPGBEGAVQWNNRLIAFASR	900	QY	1921	NWAVRTKLTPIPAASOLDLSSWVAGYSGGDIYHSLSRARPRFWFMWCLLLSVGVGIY	1980
Db	1866	FKVMSGEMPTSEDLVNLPLAILSPGALVGVVCAAILRRHVPGBEGAVQWNNRLIAFASR	1925	Db	2946	NWAVRTKLTPIPAASOLDLSSWVAGYSGGDIYHSLSRARPRFWFMWCLLLSVGVGIY	3005
QY	901	GNHVSPTHYVPSDAAARVTOILLSSLTITOLLKRLHOMINEDCSTPCSGMLRDVMDWIC	960	QY	1981	LLPNR 1985	
Db	1926	GNHVSPTHYVPSDAAARVTOILLSSLTITOLLKRLHOMINEDCSTPCSGMLRDVMDWIC	1985	Db	3006	LLPNR 3010	
QY	961	TVLDFKTWLOSKLPLRPGVPFFSCQRYKGVWRGDMIMQTTCPGQAQITGHVKNVGNMR	1020				
Db	1986	TVLDFKTWLOSKLPLRPGVPFFSCQRYKGVWRGDMIMQTTCPGQAQITGHVKNVGNMR	2045				
QY	1021	IVGPRTCNTWHTGTFPIINAYTTGCTPSPAPNYSRALMRVAEEVVEVTRVGDFFHYVTGM	1080				
Db	2046	IVGPRTCNTWHTGTFPIINAYTTGCTPSPAPNYSRALMRVAEEVVEVTRVGDFFHYVTGM	2105				

RESULT 13  
US-08-952-981A-2  
; Sequence 2, Application US/08952981A  
; Patent No. 6383768  
; GENERAL INFORMATION:  
; APPLICANT: DE FRANCESCO, Raffaele  
; APPLICANT: TOMELI, Licia  
; APPLICANT: BEHKENS, Sven-Erik

; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA  
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE  
; FILE REFERENCE: IT0002P  
; CURRENT APPLICATION NUMBER: US/08/952,981A  
; CURRENT FILING DATE: 1998-03-23  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: cDNA clone pCD (38-9.4)  
US-08-952-981A-2

Query Match		97.2%;	Score 10170;	DB 3;	Length 2201;	
Best Local Similarity		96.7%;	Pred. No. 0;			
Matches 1919;		Conservative 36;	Mismatches 30;	Indels 0;	Gaps 0;	
QY	1	MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	60			
DB	217	LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	276			
QY	61	GSKTLAGKPIITOMYTNVDQVLQKQAPPGARSITPCTCGSSDLYLVTRHADVIPVRRR	120			
DB	277	GSKTLAAPKPIITOMYTNVDQVLQKQAPPGARSITPCTCGSSDLYLVTRHADVIPVRRR	336			
QY	121	GDSRGLSPRPVSYLKGSSGGPCLCPGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180			
DB	337	GDSRGLSPRPVSYLKGSSGGPCLCPGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	396			
QY	181	RSVPFTDNSSPFAVQTPQVAHLHAPTGSKSTKVPAAAYAGQYKVLVNFNSVAATLGF	240			
DB	397	RSVPFTDNSSPFAVQTPQVAHLHAPTGSKSTKVPAAAYAGQYKVLVNFNSVAATLGF	456			
QY	241	AYMSKAHGDINIRTCVNTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDST	300			
DB	457	AYMSKAHGDINIRTCVNTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDST	516			
QY	301	ILGTVLDQAEAGARLVVLAATPPGVSVPVPHNIEVALSSTGEIPFYKAIPIETI	360			
DB	517	ILGTVLDQAEAGARLVVLAATPPGVSVPVPHNIEVALSSTGEIPFYKAIPIETI	576			
QY	361	KGRHLLIFCHSKKCKDELAALKSLGLGNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT	420			
DB	577	KGRHLLIFCHSKKCKDELAALKSLGLGNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT	636			
QY	421	GDFDSVIDCNTCVTQVDFSLDPTFTTITTPQDAVSRSQRRGTRGMRGIYRFVTPG	480			
DB	637	GDFDSVIDCNTCVTQVDFSLDPTFTTITTPQDAVSRSQRRGTRGMRGIYRFVTPG	696			
QY	481	ERPSGMFDSVILCECYDAGCAWYELTPAETSVRLRAYINTPGLPVQCQDHLEFWESVFTGL	540			
DB	697	ERPSGMFDSVILCECYDAGCAWYELTPAETSVRLRAYINTPGLPVQCQDHLEFWESVFTGL	756			
QY	541	THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDMKCLRLKPTLHGPTPL	600			
DB	757	THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDMKCLRLKPTLHGPTPL	816			
QY	601	YRLGAVQNEVTHPIPIKYIMACMSADLEVVTSTWVLVGGVLAALAYCLTSGSWIVGR	660			
DB	817	YRLGAVQNEVTHPIPIKYIMACMSADLEVVTSTWVLVGGVLAALAYCLTSGSWIVGR	876			
QY	661	IILSGKPAIIPDRVLVREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQTATKQAEA	720			
DB	877	IILSGKPAIIPDRVLVREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQTATKQAEA	936			
QY	721	AAPVVEKWRTELEAFWAKHWNFTISGTYLAGLSTLPCNPAIASLMAFTASITSLTQH	780			
DB	937	AAPVVEKWRTELEAFWAKHWNFTISGTYLAGLSTLPCNPAIASLMAFTASITSLTQH	996			
QY	781	TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDIILAGYGAGVAGALVA	840			
DB	997	TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDIILAGYGAGVAGALVA	1056			

QY	841	FKVMGEMSTEDLVNLLPAILSPGALVVGVCVCAAILRRHVGPGEGAVQMMRLIAFASR	900
DB	1057	FKVMGEMSTEDLVNLLPAILSPGALVVGVCVCAAILRRHVGPGEGAVQMMRLIAFASR	1116
QY	901	GNHVSPTHVVPESDAAARVTQILSSITITQLLKRHLQHWINEDCSTPCSGSLRDVWDWIC	960
DB	1117	GNHVSPTHVVPESDAAARVTQILSSITITQLLKRHLQHWINEDCSTPCSGSLRDVWDWIC	1176
QY	961	TVLTDPKTWLQSKLLPRLPGVPEFFSCQYKGVWBGDMQTTCCGAQITGHVXNGSMR	1020
DB	1177	TVLTDPKTWLQSKLLPRLPGVPEFFSCQYKGVWBGDMQTTCCGAQITGHVXNGSMR	1236
QY	1021	IVGPKTCNTNTHGTFPINAITYTGPCTSPAPNYSRALMRVAEEVEVTRVGDHFVVTGM	1080
DB	1237	IVGPKTCNTNTHGTFPINAITYTGPCTSPAPNYSRALMRVAEEVEVTRVGDHFVVTGM	1296
QY	1081	TTDNVYKCPQVPAPEFFTEVDGVRHRYAPACKPLLRBEVTFVLGNOYLVGSQLPCEPE	1140
DB	1297	TTDNVYKCPQVPAPEFFTEVDGVRHRYAPACKPLLRBEVTFVLGNOYLVGSQLPCEPE	1356
QY	1141	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSLKATCTTHHVSPPAD	1200
DB	1357	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSLKATCTTHHVSPPAD	1416
QY	1201	LIEANLLWRQBMGNNITRVESENKVILDSPEPQAEEDEREVSVPAILRRSRKFPRAAM	1260
DB	1417	LIEANLLWRQBMGNNITRVESENKVILDSPEPQAEEDEREVSVPAILRRSRKFPRAAM	1476
QY	1261	PIWAPDYNPPLLSWKDPDYVPVHGCPIPPAKAPPIPPRRKRTVVLSSTVSSALA	1320
DB	1477	PIWAPDYNPPLLSWKDPDYVPVHGCPIPPAKAPPIPPRRKRTVVLSSTVSSALA	1536
QY	1321	ELAKTFGSSSSAVDGTATASDPQSDGDSGDSVESYSSMPLEGEPCDPLSDGSM	1380
DB	1537	ELAKTFGSSSSAVDGTATASDPQSDGDSGDSVESYSSMPLEGEPCDPLSDGSM	1596
QY	1381	STVSEASESDVCCMSYTTGALITPCAAETKLIPINALSNSLLRHHNLVYATTSRAS	1440
DB	1597	STVSEASESDVCCMSYTTGALITPCAAETKLIPINALSNSLLRHHNLVYATTSRAS	1656
QY	1441	LROKVTDFRLQVLDHYRDLKEMKAKASTVKAALLSVEEACKLTTPHSARSKFGYAK	1500
DB	1657	LROKVTDFRLQVLDHYRDLKEMKAKASTVKAALLSVEEACKLTTPHSARSKFGYAK	1716
QY	1501	DVRNLSSKANVNHJRSVMKDLLEDTEPDTITIMAKNEVFCVQPEKGGKPARLIVFPDLG	1560
DB	1717	DVRNLSSKANVNHJRSVMKDLLEDTEPDTITIMAKNEVFCVQPEKGGKPARLIVFPDLG	1776
QY	1561	VRVCEKVALYDVVSTLTPQAVMGSSYGFQYSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	1620
DB	1777	VRVCEKVALYDVVSTLTPQAVMGSSYGFQYSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	1836
QY	1621	TVTENDIRVESIYQCCDLAPEARQATRSITERYIIGPLTNSKGQCGVRRRCRASGVL	1680
DB	1837	TVTENDIRVESIYQCCDLAPEARQATRSITERYIIGPLTNSKGQCGVRRRCRASGVL	1896
QY	1681	TSCGNTLTCYLKAAAACRAAKLODCTMLVCGDDLWVLCESAGTQDEASLRAFTTEAMTRY	1740
DB	1897	TSCGNTLTCYLKAAAACRAAKLODCTMLVCGDDLWVLCESAGTQDEASLRAFTTEAMTRY	1956
QY	1741	SAPGPPPPPEYDLELITSCSNVSVVAHDASGKRVTYLTDRDPTTPLARAWEATARTPVN	1800
DB	1957	SAPGPPPPPEYDLELITSCSNVSVVAHDASGKRVTYLTDRDPTTPLARAWEATARTPVN	2016
QY	1801	SWLGNLIMVAPTLMWARMILMTHFFSILLAQEKLKALDQCIYGCYSIEPLDLPQIIRL	1860
DB	2017	SWLGNLIMVAPTLMWARMILMTHFFSILLAQEKLKALDQCIYGCYSIEPLDLPQIIRL	2076
QY	1861	HGLSAFSLHSYSGEINRVASCLRLKGVPLRVHRHRSVRARLLSQGGAATCGKYL	1920
DB	2077	HGLSAFSLHSYSGEINRVASCLRLKGVPLRVHRHRSVRARLLSQGGAATCGKYL	2136

QY 1921 NNAVTKLKLTPIPASQDLSSWFWAGYSGGDIYHLSRPRPFWMCLLLSVGVGIY 1980  
DB 2137 NNAVTKLKLTPIPASRLDLSGFVAGYSGGDIYHLSRPRPFWMCLLLSVGVGIY 2196  
QY 1981 LLPNR 1985  
DB 2197 LLPNR 2201

RESULT 14  
US-08-324-977-36  
; Sequence 36, Application US/08324977  
; Patent No. 5747339  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto  
; APPLICANT: FURE, Isao  
; APPLICANT: MORI, Chisato  
; APPLICANT: TAKAMIZAWA, Akahisa  
; APPLICANT: YOSHIDA, Iwao  
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLend &  
; ADDRESSEE: Naughton  
; STREET: 1725 K St. N.W. Suite 1000  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/324,977  
; FILING DATE: 18-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-167466  
; FILING DATE: 25-JUN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-230921  
; FILING DATE: 31-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-305605  
; FILING DATE: 09-NOV-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/099,706  
; FILING DATE: 30-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/769,996  
; FILING DATE: 02-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/635,451  
; FILING DATE: 28-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stevens-Smith, Theresa M.  
; REGISTRATION NUMBER: 36,281  
; REFERENCE/DOCKET NUMBER: 900703D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 659-2930  
; TELEFAX: (202) 887-0357  
; TELEX: 440142  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2621 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-324-977-36

Query Match 97.2%; Score 10170; DB 1; Length 2621;

Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLGCIITSLTGRDNQVEGEVQVWSTATOSFLATCVCNGVWTVYHGA 60  
DB 637 LAPITAYSQOTRGLLGCIITSLTGRDNQVEGEVQVWSTATOSFLATCVCNGVWTVYHGA 696  
QY 61 GSKTLAGPKGITOMYTNVDQLVGMQAPPGARSITPCTCGSSDLVLVTRHADVIPVRRR 120  
DB 697 GSKTLAAPKGPITOMYTNVDQLVGMQAPPGARSITPCTCGSSDLVLVTRHADVIPVRRR 756  
QY 121 GDSRGSLLSPRPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
DB 757 GDSRGSLLSPRPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 816  
QY 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSKSTKVAAYAAQGYKVLVNLNPSVAATLFGF 240  
DB 817 RSPVFTDNSSPPAVPQSFQVAHLHAPTGSKSTKVAAYAAQGYKVLVNLNPSVAATLFGF 876  
QY 241 AYMSKAHGIDPNIRTGVRTITTCAPITYSTYTGKFLADGGCGGAYDIIICDECHSTDSTT 300  
DB 877 AYMSKAHGIDPNIRTGVRTITTCAPITYSTYTGKFLADGGCGGAYDIIICDECHSTDSTT 936  
QY 301 ILGIGTVLDOAETAGARLVVLAATATPPGVTVPHPNIEEVALSGTGEIPYGAIPETI 360  
DB 937 ILGIGTVLDOAETAGARLVVLAATATPPGVTVPHPNIEEVALSGTGEIPYGAIPETI 996  
QY 361 KGRHLIFCHSKKKCDLAAKSLGNGAVAYRGLDVSVPITSGDVVVVATDALMTGFT 420  
DB 997 RGRHLIFCHSKKKCDLAAKSLGNGAVAYRGLDVSVPITSGDVVVVATDALMTGFT 1056  
QY 421 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGMGIYRFVTPG 480  
DB 1057 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGMGIYRFVTPG 1116  
QY 481 ERPSGMPDSSVLCEDYDAGCAWYELTAPETSVBRLRAYLNTPLPVCODHLEFVESVTGL 540  
DB 1117 ERPSGMPDSSVLCEDYDAGCAWYELTAPETSVBRLRAYLNTPLPVCODHLEFVESVTGL 1176  
QY 541 THIDAHLFSLQKAGDNFPYLVAQATVCARAQAPPSDQWQKCLRLKPLTGHGTPLL 600  
DB 1177 THIDAHLFSLQKAGDNFPYLVAQATVCARAQAPPSDQWQKCLRLKPLTGHGTPLL 1236  
QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGSVTVGR 660  
DB 1237 YRLGAVQNEVTLTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGSVTVGR 1296  
QY 661 ILSGKPAIIPDREVLYREFDEMBECASHLPYIEQGMQLAEQFKQKALGILLQATKQAEA 720  
DB 1297 ILSGKPAIIPDREVLYREFDEMBECASHLPYIEQGMQLAEQFKQKALGILLQATKQAEA 1356  
QY 721 AAPVVEKWRITLFAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH 780  
DB 1357 AAPVVEKWRITLFAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH 1416  
QY 781 TLLFNILGGWVAQAOLAPPSAASAEVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
DB 1417 TLLFNILGGWVAQAOLAPPSAASAEVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1476  
QY 841 PKVMSGMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAGVWNNRLIAFASR 900  
DB 1477 PKVMSGMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAGVWNNRLIAFASR 1536  
QY 901 GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSWLRDWDWIC 960  
DB 1537 GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSWLRDWDWIC 1596  
QY 961 TVLTDFKTLQSKLLPLPGVPFFSCQSGYKVGWRGDGIMOTTCPCGAQITGHVKNQSMR 1020  
DB 1597 TVLTDFKTLQSKLLPLPGVPFFSCQSGYKVGWRGDGIMOTTCPCGAQITGHVKNQSMR 1656  
QY 1021 IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALWRVAEEYVEVTRVGDHFHYTGM 1080  
DB 1021 IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALWRVAEEYVEVTRVGDHFHYTGM 1080

Db 1657 IVGPKTCSNTWHGTFPPINAYTTGCTPTSPAPNYSRALWRVAEEYVEVTRVGDHFHYTGM 1716  
QY 1081 TTDNVKPCQVPAPPEFTVDGVRHRYAPACKPLREBEVTVLGNQYLVGSQLCPEPE 1140  
Db 1717 TTDNVKPCQVPAPPEFTVDGVRHRYAPACKPLREBEVTVLGNQYLVGSQLCPEPE 1776  
QY 1141 PDVAVLTSMLTDPSSHITATKARRLARGSPPLSSASSASQLSAPSUKATCTTHHSDPDAD 1200  
Db 1777 PDVAVLTSMLTDPSSHITATKARRLARGSPPLSSASSASQLSAPSUKATCTTHHSDPDAD 1836  
QY 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSYPAEILRRSRKPPRAM 1260  
Db 1837 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSYPAEILRRSRKPPRAM 1896  
QY 1261 PIWARDYNPELLESKWDYPPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1897 PIWARDYNPELLESKWDYPPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1956  
QY 1321 ELATKTFGSSSADVSGTATASPDQPSDDGAGSDVESYSSMPLEGEPPGDPDLSDGSW 1380  
Db 1957 ELATKTFGSSSADVSGTATALPOASDDGDKGSDVESYSSMPLEGEPPGDPDLSDGSW 2016  
QY 1381 STVSEBESADVCCSMSTYTWCALITPCAAEETKLPINALSNSLLRHNLVATTSRAS 1440  
Db 2017 STVSEBESADVCCSMSTYTWCALITPCAAEESKLPINALSNSLLRHNLVATTSRAG 2076  
QY 1441 LRQKVTFRDQLVDHVDVLEKAKASTVKALLSVEEACKLTPHSAKSKFGYGA 1500  
Db 2077 LRQKVTFRDQLVDHVDVLEKAKASTVKALLSVEEACKLTPHSAKSKFGYGA 2136  
QY 1501 DVRLSSKAVNHRSVWKDLEDTETPTDITIMAKNEVFCVQPEKGGKRPALIVFPDLG 1560  
Db 2137 DVRLSSKAVNHRSVWKDLEDTETPTDITIMAKNEVFCVQPEKGGKRPALIVFPDLG 2196  
QY 1561 VRVCEKALYDVWSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMFAIDTRCFDS 1620  
Db 2197 VRVCEKALYDVWSTLPQAVMGSSYGFQYSPQORVEFLVNTWKSNNPMGFSYDTRCFDS 2256  
QY 1621 TVTENDIRVEESIYOCCLDAPAEARQAIISLTERLYIGGPLTNSKQNGCYRRCRASGVL 1680  
Db 2257 TVTENDIRVEESIYOCCLDAPAEARQAIISLTERLYIGGPLTNSKQNGCYRRCRASGVL 2316  
QY 1681 TSCGNTLTCYLAACAAKRAKLDQCTMLVCGDDLVWICESAGTODEASLRAFTEAMTRY 1740  
Db 2317 TSCGNTLTCYLAACAAKRAKLDQCTMLVCGDDLVWICESAGTODEASLRAFTEAMTRY 2376  
QY 1741 SAPGDPKPYDLELITSCSSNVSVADHDAGKRVYLTDRPTTTLARAAMETARHTFPV 1800  
Db 2377 SAPGDPKPYDLELITSCSSNVSVADHDAGKRVYLTDRPTTTLARAAMETARHTFPV 2436  
QY 1801 SWLGNIIWYAPTLWARMILTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIRL 1860  
Db 2437 SWLGNIIWYAPTLWARMILTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIRL 2496  
QY 1861 HGLSAFSLHSYSGPEINRVASCLKGLVPLVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 2497 HGLSAFSLHSYSGPEINRVASCLKGLVPLVWRHARSVRARLLSQGGRAATCGKYL 2556  
QY 1921 NWAIVTKLTPPAAQSLDSSWFAVYGGDIYHLSLRARPFWFMCILLLSVGVGIY 1980  
Db 2557 NWAIVTKLTPPAAQSLDSSWFAVYGGDIYHLSLRARPFWFMCILLLSVGVGIY 2616  
QY 1981 LLPNR 1985  
Db 2617 LLPNR 2621

RESULT 15  
US-08-384-616-36  
; Sequence 36, Application US/08384616  
; Patent No. 5847101  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto

APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &  
ADDRESSEE: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/384,616  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
TELEX: 440142  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2621 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-384-616-36

Query Match 97.2%; Score 10170; DB 2; Length 2621;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;  
QY 1 MAPITAYSOQTRGLGCIITSLTGRDRNQVEGEVQVVSATQSFATCVNGVCTVYHGA 60  
Db 637 LAPITAYSOQTRGLGCIITSLTGRDRNQVEGEVQVVSATQSFATCVNGVCTVYHGA 696  
QY 61 GSXTLAGPKGITQMTYNVDQDLVGHQADPGARSLTPTCGSSDLVLTTHADVIPVRRR 120  
Db 697 GSXTLAGPKGITQMTYNVDQDLVGHQADPGARSLTPTCGSSDLVLTTHADVIPVRRR 756  
QY 121 GDSRGLSPRPVSYLKGSSGGLLCPSPGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 757 GDSRGLSPRPVSYLKGSSGGLLCPSPGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 816  
QY 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLINPNSVAATLGF 240  
Db 817 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLINPNSVAATLGF 876

QY	241	AYMSKAHGDIPNIRTVRTITGAPITSTYKFIADGCGSGAYDIIICDECHSTDSTT	300
Db	877	AYMSKAHGDIPNIRTVRTITGAPITSTYKFIADGCGSGAYDIIICDECHSTDSTT	936
QY	301	ILGIGTVLQDAETAGARLVVATATPPGVTVPHPNIEVALSSTGEIPFYKAIPIETI	360
Db	937	ILGIGTVLQDAETAGARLVVATATPPGVTVPHPNIEVALSSTGEIPFYKAIPIEAI	996
QY	361	KGRHILIFCHSKKKCDLAALKSLGGLNAVAYYRGDLVSVIPISGDIIVVATDALTGPT	420
Db	997	KGRHILIFCHSKKKCDLAALKSLGGLNAVAYYRGDLVSVIPISGDIIVVATDALTGPT	1056
QY	421	GFDSVIDNCTVQTVDFSLDPTFTIETTTVPQDAVSQRGRGGRGMYRVFTPG	480
Db	1057	GFDSVIDNCTVQTVDFSLDPTFTIETTTVPQDAVSQRGRGGRGMYRVFTPG	1116
QY	481	ERPSGMFDSVLCEDYDAGCAWVELTPAETSVELRAYLNTPGLPVCQDHLFEFVESFTGL	540
Db	1117	ERPSGMFDSVLCEDYDAGCAWVELTPAETSVELRAYLNTPGLPVCQDHLFEFVESFTGL	1176
QY	541	THIDAHFLSOTKQAGDNFPYIVAYQATVCARAQAPPSWDQWKCLIRLKPRTLHGPTPL	600
Db	1177	THIDAHFLSOTKQAGDNFPYIVAYQATVCARAQAPPSWDQWKCLIRLKPRTLHGPTPL	1236
QY	601	YRLGAVONEVTHPTITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSGVIVGR	660
Db	1237	YRLGAVONEVTHPTITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSGVIVGR	1296
QY	661	IILSGKPAITPDRVLYRFEDEMEECASHLPYIEQGMQLAEQFKQKALGLLQATKQAEA	720
Db	1297	IILSGKPAITPDRVLYRFEDEMEECASHLPYIEQGMQLAEQFKQKALGLLQATKQAEA	1356
QY	721	AAPVVEKWTLEAFWAKHWNFTSGIYLAGLSTLFCNPAIASLMAFTASITSLPTQH	780
Db	1357	AAPVVEKWTLEAFWAKHWNFTSGIYLAGLSTLFCNPAIASLMAFTASITSLPTQH	1416
QY	781	TLFNIIIGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYAGVAGALVA	840
Db	1417	TLFNIIIGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYAGVAGALVA	1476
QY	841	FKWMSGMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILAFASR	900
Db	1477	FKWMSGMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRILAFASR	1536
QY	901	GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSMLRDVWDWIC	960
Db	1537	GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSMLRDVWDWIC	1596
QY	961	TVLTDFTKTLQSKLLPRLPGVPPFSCQGYKGVWRGDIMQTTCPGCAQITGHVKNQSMR	1020
Db	1597	TVLTDFTKTLQSKLLPRLPGVPPFSCQGYKGVWRGDIMQTTCPGCAQITGHVKNQSMR	1656
QY	1021	IVGPRTCSTNTHGTFPPINAYTTGCTPSPAPNYGRALMRVAAEYVEVTRVGDHYVTGM	1080
Db	1657	IVGPRTCSTNTHGTFPPINAYTTGCTPSPAPNYGRALMRVAAEYVEVTRVGDHYVTGM	1716
QY	1081	TTDNVKKPCQVPAPEFFTEVDGVRHRYAPACKPLLRREEVFLVGLNQYLVGSQLPCBPE	1140
Db	1717	TTDNVKKPCQVPAPEFFTEVDGVRHRYAPACKPLLRREEVFLVGLNQYLVGSQLPCBPE	1776
QY	1141	PDVAULTSMLTDPHSITAEATAKRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD	1200
Db	1777	PDVAULTSMLTDPHSITAEATAKRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD	1836
QY	1201	LIEANLLWRQEMGNIITRVESENKVVILDSFEPLQAEEDEREVSVPABILRSKFFPRAM	1260
Db	1837	LIEANLLWRQEMGNIITRVESENKVVILDSFEPLQAEEDEREVSVPABILRSKFFPRAM	1896
QY	1261	PIWAPFDYNPPLLESKODPDYVPPVHVHGCPLPPAKAPPIPPPRRRTTVLSESTVSSALA	1320
Db	1897	PIWAPFDYNPPLLESKODPDYVPPVHVHGCPLPPAKAPPIPPPRRRTTVLSESTVSSALA	1956
QY	1321	ELATKTFGSSSESAVDSGTATASPPQSDDGAGSDVESYSNMPPILEGEPGPDLDSDGSW	1380
Db	1957	ELATKTFGSSSESAVDSGTATALPDQASDDGDKGSDVESYSNMPPILEGEPGPDLDSDGSW	2016
QY	1381	STVSEASEDVVCCSMYSYTWTCALITPCAAEETKLPINALSNLLRHHNLVATTGSRAS	1440
Db	2017	STVSEASEDVVCCSMYSYTWTCALITPCAAEESKLPINALSNLLRHHNLVATTGSRAS	2076
QY	1441	LQKQKVTFRLOVLDHHRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA	1500
Db	2077	LQKQKVTFRLOVLDHHRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA	2136
QY	1501	DYRNLSSKAVNHRSVWKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560
Db	2137	DYRNLSSKAVNHRSVWKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	2196
QY	1561	VRYCEKQALYDVVSTLPOAVNGSSYQFQSPGQVFEFLYNKAWKAKCPMGFAVDTTCFDS	1620
Db	2197	VRYCEKQALYDVVSTLPOAVNGSSYQFQSPGQVFEFLYNKAWKAKCPMGFAVDTTCFDS	2256
QY	1621	TVTENDIRVEESTYQCCDLAPPEARQAIRSLITERLYTIGGPLTNSKGQNGCYRRCRASGVL	1680
Db	2257	TVTENDIRVEESTYQCCDLAPPEARQAIRSLITERLYTIGGPLTNSKGQNGCYRRCRASGVL	2316
QY	1681	TSCGNTLTCYLKAAACRAAKLQDCTMLVCGDDLVI CESAGTORDEASLRATEAMTRY	1740
Db	2317	TSCGNTLTCYLKAAACRAAKLQDCTMLVCGDDLVI CESAGTORDEASLRATEAMTRY	2376
QY	1741	SAPPDPPKPEYDELELITSCSSNVSVAHADSGKEVYLTREDPTTPLARAAMETARHTFVN	1800
Db	2377	SAPPDPPKPEYDELELITSCSSNVSVAHADSGKEVYLTREDPTTPLARAAMETARHTFVN	2436
QY	1801	SWLGNIIIMYAPTILWARMILMTHFFSILLAQEQLKALDCOIYGACYSIEPLDLQIIFORL	1860
Db	2437	SWLGNIIIMYAPTILWARMILMTHFFSILLAQEQLKALDCOIYGACYSIEPLDLQIIFORL	2496
QY	1861	HGLSAFSLHSYSFGEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRATCGKYL	1920
Db	2497	HGLSAFSLHSYSFGEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRATCGKYL	2556
QY	1921	NWAVRKLKLTPTPAASQDLSSWFVAGYSGGDIYHLSRARPRFWMCLILLSVGVGIY	1980
Db	2557	NWAVRKLKLTPTPAASQDLSSWFVAGYSGGDIYHLSRARPRFWMCLILLSVGVGIY	2616
QY	1981	LLPNR 1985	
Db	2617	LLPNR 2621	

Search completed: December 8, 2004, 12:33:52  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:25:07 ; Search time 129 Seconds

(without alignments)  
5496.116 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQOTRGLGCIIT.....FMWCLLLSVGVIGVILLPNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10465	100.0	1985	16	US-10-639-150-2
2	10462	100.0	3010	15	US-10-467-000-1
3	10459	99.9	1985	14	US-10-259-275-42
4	10442	99.8	2201	13	US-10-029-907-3
5	10442	99.8	2201	14	US-10-309-561-3
6	10442	99.8	2201	16	US-10-789-355-3
7	10442	99.8	2201	17	US-10-686-835-3
8	10170	97.2	2201	13	US-10-085-476-2
9	10157	97.1	3010	16	US-10-333-449A-34
10	9710	92.8	2985	14	US-10-259-275-40
11	9454.5	90.3	3011	15	US-10-296-734-406
12	9434.5	90.2	3011	9	US-09-742-659-4
13	9430.5	90.1	3011	10	US-09-891-894-3
					Sequence 2, Appli
					Sequence 1, Appli
					Sequence 42, Appl
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 34, Appl
					Sequence 40, Appl
					Sequence 406, App
					Sequence 4, Appli
					Sequence 3, Appli

14	9430.5	90.1	3011	14	US-10-184-150-3	Sequence 3, Appli
15	9430.5	90.1	3011	14	US-10-328-997-3	Sequence 3, Appli
16	9430.5	90.1	3012	9	US-09-238-076-2	Sequence 2, Appli
17	9430.5	90.1	3012	10	US-09-995-937-2	Sequence 2, Appli
18	9430.5	90.1	3012	10	US-09-917-583-2	Sequence 2, Appli
19	9418.5	90.0	3011	9	US-09-952-572-9	Sequence 20, Appl
20	9418.5	90.0	3011	9	US-09-747-419-20	Sequence 9, Appli
21	9418.5	90.0	3011	14	US-10-259-275-20	Sequence 20, Appl
22	9418.5	90.0	3011	15	US-10-189-359-14	Sequence 14, Appl
23	9417.5	90.0	3011	9	US-09-916-359-2	Sequence 2, Appli
24	9417.5	90.0	3011	16	US-10-445-724-2	Sequence 2, Appli
25	9371.5	89.6	3011	9	US-09-238-076-20	Sequence 20, Appl
26	9371.5	89.6	3011	10	US-09-995-937-20	Sequence 20, Appl
27	9371.5	89.6	3011	10	US-09-917-563-20	Sequence 20, Appl
28	9324.5	89.1	3011	14	US-10-232-643-6	Sequence 6, Appli
29	9266.5	88.5	3011	9	US-09-929-955-1	Sequence 1, Appli
30	9266.5	88.5	3011	13	US-10-104-966-1	Sequence 1, Appli
31	9266.5	88.5	3011	15	US-10-719-619-1	Sequence 1, Appli
32	9266.5	88.5	3011	17	US-10-817-591-1	Sequence 1, Appli
33	8799.5	84.1	2894	9	US-09-941-611-23	Sequence 23, Appl
34	8799.5	84.1	2894	14	US-10-044-995-23	Sequence 23, Appl
35	7638.5	73.0	2940	14	US-10-226-629A-13	Sequence 13, Appl
36	7168	68.5	1736	14	US-10-328-127-2	Sequence 2, Appli
37	7168	68.5	1736	14	US-10-328-206-2	Sequence 2, Appli
38	4823.5	46.1	2865	9	US-09-742-659-6	Sequence 6, Appli
39	4725.5	45.2	1040	9	US-09-929-955-9	Sequence 9, Appli
40	4725.5	45.2	1040	13	US-10-104-966-9	Sequence 9, Appli
41	4725.5	45.2	1040	15	US-10-719-619-9	Sequence 9, Appli
42	4725.5	45.2	1040	17	US-10-817-591-9	Sequence 9, Appli
43	4443	42.5	2307	10	US-09-919-301-2	Sequence 2, Appli
44	4443	42.5	2307	14	US-10-191-966-2	Sequence 2, Appli
45	4440	42.4	2307	10	US-09-919-301-9	Sequence 9, Appli

#### ALIGNMENTS

RESULT 1  
US-10-639-150-2  
; Sequence 2, Application US/10639150  
; Publication No. US20040121975A1  
; GENERAL INFORMATION:  
; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY  
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS  
; FILE REFERENCE: D0224 NP  
; CURRENT APPLICATION NUMBER: US/10/639,150  
; CURRENT FILING DATE: 2003-08-12  
; PRIOR APPLICATION NUMBER: US 60/402,661  
; PRIOR FILING DATE: 2002-08-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: HCV Replicon  
; US-10-639-150-2

Query Match	100.0%;	Score	10465;	DB	16;	Length	1985;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1985;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAPITAYSQOTRGLGCIITSLTGRDRNOVEGEVQVQVSTATQSFATCVNGVCMVTVHGA	60				
Db	1	MAPITAYSQOTRGLGCIITSLTGRDRNOVEGEVQVQVSTATQSFATCVNGVCMVTVHGA	60				
Qy	61	GSKTLGAPKGPITQMTYNVDQDLVGWQAPPGARSLLPCTCGSSDLYLVTRHADVIPVRRR	120				
Db	61	GSKTLGAPKGPITQMTYNVDQDLVGWQAPPGARSLLPCTCGSSDLYLVTRHADVIPVRRR	120				
Qy	121	GDGRGSLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180				
Db	121	GDGRGSLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180				



```

Db 121 GDSRGLSPRVSVYLKSGSGGLLPCPSGHAVGI FRAAVCTRGVAKAVDFVPVESMETTM 180
QY 181 RSPVFTDNSSPPAVPQTQVAVHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240
Db 181 RSPVFTDNSSPPAVPQTQVAVHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240
QY 241 AYSKAGHDNIRTOVRTITGAPITYSTYKFLADGCGSGAYDIIICDECHSDSTT 300
Db 241 AYSKAGHDNIRTOVRTITGAPITYSTYKFLADGCGSGAYDIIICDECHSDSTT 300
QY 301 ILGIGVLDQAEATAGARLVVLAATAPPGSVTVPHPNIEVALSGTGEIPYKAIPIETI 360
Db 301 ILGIGVLDQAEATAGARLVVLAATAPPGSVTVPHPNIEVALSGTGEIPYKAIPIETI 360
QY 361 KGRHLIFCHSKKKKDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 420
Db 361 KGRHLIFCHSKKKKDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 420
QY 421 GDFSVIDCNCVTQTVDFSLDPTFTIETTVPQDAVSRSORGRGGRMGYIRFVTPG 480
Db 421 GDFSVIDCNCVTQTVDFSLDPTFTIETTVPQDAVSRSORGRGGRMGYIRFVTPG 480
QY 481 ERPSGMFDSVILCECYDAGCAWELTPAETSRLRAYLNTPLPVCQDHLFEFVESVFTGL 540
Db 481 ERPSGMFDSVILCECYDAGCAWELTPAETSRLRAYLNTPLPVCQDHLFEFVESVFTGL 540
QY 541 THIDAHFLSQTQAGDNFPVLYAYQATVCARAQAPPPSWDMKCLIRLXPTLHGTPPLL 600
Db 541 THIDAHFLSQTQAGDNFPVLYAYQATVCARAQAPPPSWDMKCLIRLXPTLHGTPPLL 600
QY 601 YRLGAVONEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVTVGR 660
Db 601 YRLGAVONEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVTVGR 660
QY 661 IILSGPAILPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLLATKQAEA 720
Db 661 IILSGPAILPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLLATKQAEA 720
QY 721 AAPVVEKWTLEAFWAKHWNFTSGTQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780
Db 721 AAPVVEKWTLEAFWAKHWNFTSGTQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780
QY 781 TLLFNILGGWVAQALAPPSAASAVGAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 840
Db 781 TLLFNILGGWVAQALAPPSAASAVGAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 840
QY 841 FKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRLIAFASR 900
Db 841 FKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRLIAFASR 900
QY 901 GNVSPHYVPESDAAARVTOILSSLIITQLLKLHOWINEDCSTPGSGWLRDWDWIC 960
Db 901 GNVSPHYVPESDAAARVTOILSSLIITQLLKLHOWINEDCSTPGSGWLRDWDWIC 960
QY 961 TVLDFKTLQSKLLPRLPGVPFFSCORGKYGWVRGIMOTTCPCQAQITGHVKGSMR 1020
Db 961 TVLDFKTLQSKLLPRLPGVPFFSCORGKYGWVRGIMOTTCPCQAQITGHVKGSMR 1020
QY 1021 IVGPRTCSNTHWGTFFINAYTTGPTSPAPNYGRALWRVAEEYVEVTRVGDHYVTGM 1080
Db 1021 IVGPRTCSNTHWGTFFINAYTTGPTSPAPNYGRALWRVAEEYVEVTRVGDHYVTGM 1080
QY 1081 TTDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLAREVTVFLGQNYLNGSOLPCEPE 1140
Db 1081 TTDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLAREVTVFLGQNYLNGSOLPCEPE 1140
QY 1141 PDVAVLTSMLTDPGHITAEAKRLARGSPSLASSASQLSAPSKATCTTRHDSPAD 1200
Db 1141 PDVAVLTSMLTDPGHITAEAKRLARGSPSLASSASQLSAPSKATCTTRHDSPAD 1200
QY 1201 LIEANLLWRQBMGNITRVSENKVVILDSFEPLQAEDEREVSPABILRRSRKFFPRAM 1260
Db 1201 LIEANLLWRQBMGNITRVSENKVVILDSFEPLQAEDEREVSPABILRRSRKFFPRAM 1260

```

```

1261 PIWARPDPNPPLESWKDPDYVPPVVGCPPLPAKAPPIPPRRKRTTVLSESTVSALA 1320
1261 PIWARPDPNPPLESWKDPDYVPPVVGCPPLPAKAPPIPPRRKRTTVLSESTVSALA 1320
1321 ELATKTFGSSSSAVDSGTATASPDQSDDGAGSDVESYSSMPPLLEGPGDPLSDGWS 1380
1321 ELATKTFGSSSSAVDSGTATASPDQSDDGAGSDVESYSSMPPLLEGPGDPLSDGWS 1380
1381 STVSEBASDVVCCSMSTYTWGALITPCAAETKLPINALSNLSLRHNLVYATTGRSAS 1440
1381 STVSEBASDVVCCSMSTYTWGALITPCAAETKLPINALSNLSLRHNLVYATTGRSAS 1440
1441 LRQKVTYFRLQVLDHVDLKVEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500
1441 LRQKVTYFRLQVLDHVDLKVEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500
1501 DVNLSKAVNHIRSVMKOLLEDTETPTTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560
1501 DVNLSKAVNHIRSVMKOLLEDTETPTTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560
1561 VRVCEKALYDVVSTLTPQAVMGSSYGFQYSPGQVFLVNAWKAKKCPMGFAYDTRCFDS 1620
1561 VRVCEKALYDVVSTLTPQAVMGSSYGFQYSPGQVFLVNAWKAKKCPMGFAYDTRCFDS 1620
1621 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGONCGYRRCRASGVL 1680
1621 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGONCGYRRCRASGVL 1680
1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEDEASLRAFTEAMTRY 1740
1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEDEASLRAFTEAMTRY 1740
1741 SAPGDPPEPEYDELEITSCSSNVSVAHADSGKRVYVLTDRPTPLARAWEATARHTPVN 1800
1741 SAPGDPPEPEYDELEITSCSSNVSVAHADSGKRVYVLTDRPTPLARAWEATARHTPVN 1800
1801 SWLGNIMYAPTILWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRL 1860
1801 SWLGNIMYAPTILWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRL 1860
1861 HGLSAFSLHSYSGEINRVASCLRLKGVPLRVRHARSVRARLLSQGGRATCGKYL 1920
1861 HGLSAFSLHSYSGEINRVASCLRLKGVPLRVRHARSVRARLLSQGGRATCGKYL 1920
1921 NWAVRTKLTPTPAASQDLSSWVAGYSGGDIYHLSRARPRFWMCLLLSVGVGIY 1980
1921 NWAVRTKLTPTPAASQDLSSWVAGYSGGDIYHLSRARPRFWMCLLLSVGVGIY 1980
1981 LLPNR 1985
1981 LLPNR 1985

```

```

RESULT 2
US-10-467-000-1
; Sequence 1, Application US/10467000
; Publication No. US20040067486A1
; GENERAL INFORMATION:
; APPLICANT: De Francesco, Raffaele
; APPLICANT: Migliaccio, Giovanni
; APPLICANT: Paonessa, Giacomo
; TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON
; TITLE OF INVENTION: ENHANCED CELLS
; FILE REFERENCE: ITR0003P
; CURRENT APPLICATION NUMBER: US/10/467,000
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: PCT/EP02/00526
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/263,479
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 4.0

```

; SEQ ID NO 1									
; LENGTH: 3010									
; TYPE: PRT									
; ORGANISM: Con 1 HCV isolate nucleic acid									
US-10-467-000-1									
Query Match									
Best Local Similarity 100.0%; Score 10462; DB 15; Length 3010;									
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQFLATCVNGVCMVYHGA	60						
DB	1026	LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQFLATCVNGVCMVYHGA	1085						
QY	61	GSKTLAGPKGPIQMYTNVDQDLVQWQAPPQARSLLTPTCTCGSSDLYLVTRHADVIPVRRR	120						
DB	1086	GSKTLAGPKGPIQMYTNVDQDLVQWQAPPQARSLLTPTCTCGSSDLYLVTRHADVIPVRRR	1145						
QY	121	GDSRGSLLSPRPVSYLKGSSGGPILLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180						
DB	1146	GDSRGSLLSPRPVSYLKGSSGGPILLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	1205						
QY	181	RSPVFTDNSSPPAVPQTQFVAHLHAPTGSKSTKVPAAYAAQGYKVLVNFPSVAATLFGF	240						
DB	1206	RSPVFTDNSSPPAVPQTQFVAHLHAPTGSKSTKVPAAYAAQGYKVLVNFPSVAATLFGF	1265						
QY	241	AYMSKAHGIDENIRGTITTTGAPITTYSTYKFLADGCGSGGAYDIIICDECHSTDSTT	300						
DB	1266	AYMSKAHGIDENIRGTITTTGAPITTYSTYKFLADGCGSGGAYDIIICDECHSTDSTT	1325						
QY	301	ILGIGTVLDAQETAGARLVWLATATPPGVSIVPHNIEEVALSSTGEIPFGKAIPIETI	360						
DB	1326	ILGIGTVLDAQETAGARLVWLATATPPGVSIVPHNIEEVALSSTGEIPFGKAIPIETI	1385						
QY	361	KGRHLLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIWATDALMTGFT	420						
DB	1386	KGRHLLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIWATDALMTGFT	1445						
QY	421	GDFDSVDCNTCVTQVDFSLDPTFTTITTTVPQDAVSRQRGRGTGRGRGIYRFVTPG	480						
DB	1446	GDFDSVDCNTCVTQVDFSLDPTFTTITTTVPQDAVSRQRGRGTGRGRGIYRFVTPG	1505						
QY	481	ERPSCMFDSVLCBCEYDAGCAWYELTPAETSVRLRAYLINTPLPVCODHLEFWESVFTGL	540						
DB	1506	ERPSCMFDSVLCBCEYDAGCAWYELTPAETSVRLRAYLINTPLPVCODHLEFWESVFTGL	1565						
QY	541	THIDAHFLSOTKQAGDNFPYLVAQATVCARAQAPPPSWDQMKCLIRLKPTELHGPTPL	600						
DB	1566	THIDAHFLSOTKQAGDNFPYLVAQATVCARAQAPPPSWDQMKCLIRLKPTELHGPTPL	1625						
QY	601	YRLGAVONEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAYCLTTGVSIVVGR	660						
DB	1626	YRLGAVONEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAYCLTTGVSIVVGR	1685						
QY	661	IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQGMQLABEQFKOKAIGLLQATKQAEA	720						
DB	1686	IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQGMQLABEQFKOKAIGLLQATKQAEA	1745						
QY	721	AAPVVEKWRITLEAFWAKHMWNFTSGIQLAGLSTLPGNPAIASIMAFATASITSPLTQH	780						
DB	1746	AAPVVEKWRITLEAFWAKHMWNFTSGIQLAGLSTLPGNPAIASIMAFATASITSPLTQH	1805						
QY	781	TLLENILGGWAAQALAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA	840						
DB	1806	TLLENILGGWAAQALAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA	1865						
QY	841	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRIIAFASR	900						
DB	1866	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRIIAFASR	1925						
QY	901	GNHVSPTHYPESDAAARVTOILSSLTITQLLKRHLQWINEDCSTPCSGSWLRDWDWIC	960						
DB	1926	GNHVSPTHYPESDAAARVTOILSSLTITQLLKRHLQWINEDCSTPCSGSWLRDWDWIC	1985						

RESULT 3  
 US-10-259-275-42  
 ; Sequence 42, Application US/10259275  
 ; Publication No. US20030125541A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemon, Stanley M.  
 ; APPLICANT: Y1, Minkyung  
 ; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE  
 ; FILE REFERENCE: 265.0007 0120  
 ; CURRENT APPLICATION NUMBER: US/10/259,275  
 ; CURRENT FILING DATE: 2003-01-13  
 ; PRIOR APPLICATION NUMBER: US 60/171,909  
 ; PRIOR FILING DATE: 1999-12-23  
 ; PRIOR APPLICATION NUMBER: US 09/747,419  
 ; PRIOR FILING DATE: 2000-12-23  
 ; PRIOR APPLICATION NUMBER: US 60/325,236  
 ; PRIOR FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: US 60/338,123  
 ; PRIOR FILING DATE: 2001-11-13  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 42  
 ; LENGTH: 1985  
 ; TYPE: PRN  
 ; ORGANISM: ARTIFICIAL  
 ; FEATURE:  
 ; OTHER INFORMATION: amino acid sequence encoded by the nucleotides 2119-8073 of  
 ; OTHER INFORMATION: SEQ ID NO:41  
 US-10-259-275-42

Query Match 99.9%; Score 10459; DB 14; Length 1985;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQOTGGLGCIITSLGRDRNOVEGEVQVQVSTATQSFATCVNGVCWTVVHGA 60  
 DB 1 MAPITAYSQOTGGLGCIITSLGRDRNOVEGEVQVQVSTATQSFATCVNGVCWTVVHGA 60  
 QY 61 GSKTLAGKPGFITOMYTNVDQDLVGVQAPPGARSILTECTGSSDLVLTWRHADVPVRR 120  
 DB 61 GSKTLAGKPGFITOMYTNVDQDLVGVQAPPGARSILTECTGSSDLVLTWRHADVPVRR 120  
 QY 121 GDSRGLSPRVSYLKSSGGPCLCPSHAVGIFRAAVCTRGVAKAVDPVPMESMETM 180  
 DB 121 GDSRGLSPRVSYLKSSGGPCLCPSHAVGIFRAAVCTRGVAKAVDPVPMESMETM 180  
 QY 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQYKVLVLPNSVAATLFG 240  
 DB 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQYKVLVLPNSVAATLFG 240  
 QY 241 AYMSKAHGIDPNIRGTVITTTGAPITVSTYKFLADGGCGGGAYDIIICDECHSTDSTT 300  
 DB 241 AYMSKAHGIDPNIRGTVITTTGAPITVSTYKFLADGGCGGGAYDIIICDECHSTDSTT 300  
 QY 301 ILGIGTLDQAEAGARLWLATATPGSVTPVPHNIEVALSTGSIPIFYKGAIPETI 360  
 DB 301 ILGIGTLDQAEAGARLWLATATPGSVTPVPHNIEVALSTGSIPIFYKGAIPETI 360  
 QY 361 KGGRHILFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIWVATDALMTGFT 420  
 DB 361 KGGRHILFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIWVATDALMTGFT 420  
 QY 421 GDFDSVIDCNTCVITQVDFSLDPTFTTETTTVPQDAVRSORRGTRGRMGIRFVFTPG 480  
 DB 421 GDFDSVIDCNTCVITQVDFSLDPTFTTETTTVPQDAVRSORRGTRGRMGIRFVFTPG 480  
 QY 481 BRPSGMDSSVLCBCYDAGCAWYELTPAETSURLRAYLNTPLGVCQDHLFEWESVFTGL 540  
 DB 481 BRPSGMDSSVLCBCYDAGCAWYELTPAETSURLRAYLNTPLGVCQDHLFEWESVFTGL 540  
 QY 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDMWKCLIRLKTPLHGPTLL 600

DB 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDMWKCLIRLKTPLHGPTLL 600  
 QY 601 YRLGAVQNEVTTHTPITTKYIMACMSADLEWVSTWLVGVLAALAAAYCLTTCGSSVIVGR 660  
 DB 601 YRLGAVQNEVTTHTPITTKYIMACMSADLEWVSTWLVGVLAALAAAYCLTTCGSSVIVGR 660  
 QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLYPIEQMQLAEQFKQKAIQLOTTATQAE 720  
 DB 661 IILSGKPAIIPDREVLYREFDEMEECASHLYPIEQMQLAEQFKQKAIQLOTTATQAE 720  
 QY 721 AAPVVEKWTLEAFWAKEMWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITPTTQH 780  
 DB 721 AAPVVEKWTLEAFWAKEMWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITPTTQH 780  
 QY 781 TLLFNILGGWAAQAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGVAGVAGALVA 840  
 DB 781 TLLFNILGGWAAQAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGVAGVAGALVA 840  
 QY 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFAS 900  
 DB 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFAS 900  
 QY 901 GNVSPHYVPESDAARVTQILSLTITQLLRLHQMINEDCSTPCSGSLRDVMDWIC 960  
 DB 901 GNVSPHYVPESDAARVTQILSLTITQLLRLHQMINEDCSTPCSGSLRDVMDWIC 960  
 QY 961 TVLTDFKTLQSKLLPRLPGVFPFCQGVKGVWRGDMQITTCPCGAQITGHVNGSMR 1020  
 DB 961 TVLTDFKTLQSKLLPRLPGVFPFCQGVKGVWRGDMQITTCPCGAQITGHVNGSMR 1020  
 QY 1021 IVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALWRVAEEYVEVTVGVFHYVTGM 1080  
 DB 1021 IVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALWRVAEEYVEVTVGVFHYVTGM 1080  
 QY 1081 TTDNVKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEVTVFLVGLNQYLVGSLPCEPE 1140  
 DB 1081 TTDNVKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEVTVFLVGLNQYLVGSLPCEPE 1140  
 QY 1141 PDVAVLTSMLTDSHITAEAKERLARGSPSSASSASOLASPSLKATCTTRHSDSPAD 1200  
 DB 1141 PDVAVLTSMLTDSHITAEAKERLARGSPSSASSASOLASPSLKATCTTRHSDSPAD 1200  
 QY 1201 LIEANILWRQEMGNITRVESENKVVILDSFEPLQAEDEEREVSVPAAEILRRSRKFFRAM 1260  
 DB 1201 LIEANILWRQEMGNITRVESENKVVILDSFEPLQAEDEEREVSVPAAEILRRSRKFFRAM 1260  
 QY 1261 PIWARPYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVIVSESTVSALA 1320  
 DB 1261 PIWARPYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVIVSESTVSALA 1320  
 QY 1321 ELATKTFGSSSAVDSGTATASDPQSDGSDGSDVESYSSMPLEGEPCDPLSDGSW 1380  
 DB 1321 ELATKTFGSSSAVDSGTATASDPQSDGSDGSDVESYSSMPLEGEPCDPLSDGSW 1380  
 QY 1381 STVSEEAESDVCCSMSTWTGALITPCAAEETKLIPINALSNLSLRHNLVYATTSAS 1440  
 DB 1381 STVSEEAESDVCCSMSTWTGALITPCAAEETKLIPINALSNLSLRHNLVYATTSAS 1440  
 QY 1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVAKALLSVEEACKLTPHARSFGYGA 1500  
 DB 1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVAKALLSVEEACKLTPHARSFGYGA 1500  
 QY 1501 DVRLNSSKAVNHRSVWKDLLEDTPIDTITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 DB 1501 DVRLNSSKAVNHRSVWKDLLEDTPIDTITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 QY 1561 VRVCEKVALXDVSTLPAQVMGSSYGFQYSGQVREFLVNAWAKKCPMGFAYDTRCDS 1620  
 DB 1561 VRVCEKVALXDVSTLPAQVMGSSYGFQYSGQVREFLVNAWAKKCPMGFAYDTRCDS 1620  
 QY 1621 TVTENDIRVEESIYQCCDLAPARQATRSILTERLYIGGLTNSKGNCGYRCRASGVL 1680  
 DB 1621 TVTENDIRVEESIYQCCDLAPARQATRSILTERLYIGGLTNSKGNCGYRCRASGVL 1680

QY	1681	TSCGNTLTCLYKAAACRAAKLODCTMLVCGDDLVVICESACTOEDASLRAFTAMTRY	1740
Db	1681	TSCGNTLTCLYKAAACRAAKLODCTMLVCGDDLVVICESACTOEDASLRAFTAMTRY	1740
QY	1741	SAPPDPPKPEYDLELITSCSSNVSAHDASGRVYVLTFRDPTTPLARAAMETARHTPVN	1800
Db	1741	SAPPDPPKPEYDLELITSCSSNVSAHDASGRVYVLTFRDPTTPLARAAMETARHTPVN	1800
QY	1801	SWLGNIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIQR	1860
Db	1801	SWLGNIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIQR	1860
QY	1861	HGLSAFSLHSYSGEINRVASCLRKLGVPLRVWRHARSVRARLLSQGGRAATCGKYL	1920
Db	1861	HGLSAFSLHSYSGEINRVASCLRKLGVPLRVWRHARSVRARLLSQGGRAATCGKYL	1920
QY	1921	NWAVRTKLKLTPIPAASQLDLSSWFWAGYSGGDIYHSLSRARPRFWMCLLLLSVGVI	1980
Db	1921	NWAVRTKLKLTPIPAASQLDLSSWFWAGYSGGDIYHSLSRARPRFWMCLLLLSVGVI	1980
QY	1981	LLPNR 1985	
Db	1981	LLPNR 1985	
RESULT 4			
US-10-029-907-3			
; Sequence 3, Application US/10029907			
; Publication No. US20020142350A1			
; GENERAL INFORMATION:			
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.			
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM			
; TITLE OF INVENTION: HEPATITIS C VIRUS			
; FILE REFERENCE: 13/083			
; CURRENT APPLICATION NUMBER: US/10/029,907			
; CURRENT FILING DATE: 2001-12-21			
; PRIOR APPLICATION NUMBER: 60/257,857			
; PRIOR FILING DATE: 2000-12-22			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 2201			
; TYPE: PRT			
; ORGANISM: HCV			
; FEATURE:			
; NAME/KEY: VARIANT			
; LOCATION: 882			
; OTHER INFORMATION: Xaa is Lys or Arg			
; NAME/KEY: VARIANT			
; LOCATION: 1489			
; OTHER INFORMATION: Xaa is Leu			
US-10-029-907-3			
Query Match 99.8%; Score 10442; DB 13; Length 2201;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;			
QY	1	MAPTAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	60
Db	217	LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	276
QY	61	GSKTLAGPKGITOMYTNVDDLVGMQAPPGARSILTCTCGSSDLYLVTRHADVIPVRRR	120
Db	277	GSKTLAGPKGITOMYTNVDDLVGMQAPPGARSILTCTCGSSDLYLVTRHADVIPVRRR	336
QY	121	GDSRGSLLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM	180
Db	337	GDSRGSLLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM	396
QY	181	RSPVFTDNSPPAVPQTQVAHLHAPTSGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFG	240
Db	397	RSPVFTDNSPPAVPQTQVAHLHAPTSGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFG	456
QY	241	AYMSKAHGIDPNIRTVGRTITTTGAPITTYSTYKFLADGGCGSGAYDIIICDECHSDSTT	300
Db	457	AYMSKAHGIDPNIRTVGRTITTTGAPITTYSTYKFLADGGCGSGAYDIIICDECHSDSTT	516
QY	301	ILIGICTVLDOAETAGARLVVLTATATPGSVTVPHNIEVALSSGIEIPFYKAIPIETI	360
Db	517	ILIGICTVLDOAETAGARLVVLTATATPGSVTVPHNIEVALSSGIEIPFYKAIPIETI	576
QY	361	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGT	420
Db	577	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGT	636
QY	421	GFDSVIDNCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGRMGIYFVTPG	480
Db	637	GFDSVIDNCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGRMGIYFVTPG	696
QY	481	ERPSGMFSSVLCCEYDAGCAWYELTPAETSVRLRAYLNTPLGPVQDHLFEWESVFTGL	540
Db	697	ERPSGMFSSVLCCEYDAGCAWYELTPAETSVRLRAYLNTPLGPVQDHLFEWESVFTGL	756
QY	541	THIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKLPHGPTPLL	600
Db	757	THIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKLPHGPTPLL	816
QY	601	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	660
Db	817	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	876
QY	661	IILSGKPAIIPDREVLVREDFEDEMEECASHLYPIEQMQLAEQFKQKATGLIQTAKQABA	720
Db	877	IILSGKPAIIPDREVLVREDFEDEMEECASHLYPIEQMQLAEQFKQKATGLIQTAKQABA	936
QY	721	AAPVVESKWRLEAFWAKHWNFIISGLOLAGSLTPCNPAIASLMAFTASITSLTTTQH	780
Db	937	AAPVVESKWRLEAFWAKHWNFIISGLOLAGSLTPCNPAIASLMAFTASITSLTTTQH	996
QY	781	TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYGAGVAGALVA	840
Db	997	TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYGAGVAGALVA	1056
QY	841	FKVMSGEMPSTEDLVNLLPALISPCALVGVVCAAILRRHVGPGEAVQWNNRIAFASR	900
Db	1057	FKVMSGEMPSTEDLVNLLPALISPCALVGVVCAAILRRHVGPGEAVQWNNRIAFASR	1116
QY	901	GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDECTPCSGSLRVDWDMIC	960
Db	1117	GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDECTPCSGSLRVDWDMIC	1176
QY	961	TVLTDFTWLOSULLPRLPGVPPFCQGYGVWRGDGIMQTTCCGAAQITGHVKNCSMR	1020
Db	1177	TVLTDFTWLOSULLPRLPGVPPFCQGYGVWRGDGIMQTTCCGAAQITGHVKNCSMR	1236
QY	1021	IVGPRTCSTNWHGTFPINAYTTGCTSPAPNYSRALWRVAAEEYVEVTRVGDHYVTGM	1080
Db	1237	IVGPRTCSTNWHGTFPINAYTTGCTSPAPNYSRALWRVAAEEYVEVTRVGDHYVTGM	1296
QY	1081	TTDNVKKCPQVPAPEFFTEVDGVRHLRYAPACKPILLREEVTFVLGNQYLVGSLPCEPE	1140
Db	1297	TTDNVKKCPQVPAPEFFTEVDGVRHLRYAPACKPILLREEVTFVLGNQYLVGSLPCEPE	1356
QY	1141	PDVAULTSMLTDPESHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTTRHDSPAD	1200
Db	1357	PDVAULTSMLTDPESHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTTRHDSPAD	1416
QY	1201	LIEANLLWROEMGNITRVESENKVVILDSFEPLQAEDEREVSVPAEIIIRRSKFPFRAM	1260
Db	1417	LIEANLLWROEMGNITRVESENKVVILDSFEPLQAEDEREVSVPAEIIIRRSKFPFRAM	1476
QY	1261	PIWARPDYNNPPLLESWKDPDYVPVVGHCPLPPAKAPPIPPRRKRRTVILSESVSSALA	1320
Db	1477	PIWARPDYNNPPLLESWKDPDYVPVVGHCPLPPAKAPPIPPRRKRRTVILSESVSSALA	1536

```

1321 ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSSMPLEGEPPDLDGSGW 1380
1537 ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSSMPLEGEPPDLDGSGW 1596
1381 STVSEASESDVCCMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSSAS 1440
1597 STVSEASESDVCCMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSSAS 1656
1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKALLSVEEAKLTPPHSARKFGYGA 1500
1657 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKALLSVEEAKLTPPHSARKFGYGA 1716
1501 DVENLSKAVNHTRSVWVKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKEARLIVPDLG 1560
1717 DVENLSKAVNHTRSVWVKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKEARLIVPDLG 1776
1561 VRYCEKXALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAVDTTRCFDS 1620
1777 VRYCEKXALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAVDTTRCFDS 1836
1621 TVTENDIRVESYQCCDLAPEARQAIRSLTERLIYIGGPLTNSKGQNGYRRRCRAGVLT 1680
1837 TVTENDIRVESYQCCDLAPEARQAIRSLTERLIYIGGPLTNSKGQNGYRRRCRAGVLT 1896
1681 TSCGNTLT CYLKAACRAAKLOCTMLVCGDDLVWICESAGTOEDEASLRATEAMTRY 1740
1897 TSCGNTLT CYLKAACRAAKLOCTMLVCGDDLVWICESAGTOEDEASLRATEAMTRY 1956
1741 SAPPGDPKPEYDILELITSCSSNVSAVHADSGKVVYLTDPPTPLARAAMETARHTPVN 1800
1957 SAPPGDPKPEYDILELITSCSSNVSAVHADSGKVVYLTDPPTPLARAAMETARHTPVN 2016
1801 SWLGNITMYAPTLMWARMILMTHFFSILLAOBLEKALDCQIYACYSIEPLDLPOIIOQL 1860
2017 SWLGNITMYAPTLMWARMILMTHFFSILLAOBLEKALDCQIYACYSIEPLDLPOIIOQL 2076
1861 HGLSASFSLHSYSGEINRVASCLKLGVPPLRVWHRARSVRARLLSQGGRAATCGKLYF 1920
2077 HGLSASFSLHSYSGEINRVASCLKLGVPPLRVWHRARSVRARLLSQGGRAATCGKLYF 2136
1921 NWAVRTKLKLTPIPAASQDLSSWFVAGYSGGDIYHLSLRARPFWFWMCLLLSVGVGIY 1980
2137 NWAVRTKLKLTPIPAASQDLSSWFVAGYSGGDIYHLSLRARPFWFWMCLLLSVGVGIY 2196
1981 LLPNR 1985
2197 LLPNR 2201

```

```

RESULT 5
US-10-309-561-3
; Sequence 3, Application US/10309561
; Publication No. US20030148348A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10309,561
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 882

```

```

; OTHER INFORMATION: Xaa is Lys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1489
; OTHER INFORMATION: Xaa is Leu
US-10-309-561-3

Query Match          99.8%; Score 10442; DB 14; Length 2201;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPIATYSQOTRGLLCIITITSLTGRDRNQVEGVQVSTATQSFLLATCVNGVCWTVVHGA 60
DB 217 LAPIATYSQOTRGLLCIITITSLTGRDRNQVEGVQVSTATQSFLLATCVNGVCWTVVHGA 276
QY 61 GSTLAGPKGPIITQMTNVDDLVGWQAPPGASRLTPTCTGSSDDLVLVTHADVIPVRRR 120
DB 277 GSTLAGPKGPIITQMTNVDDLVGWQAPPGASRLTPTCTGSSDDLVLVTHADVIPVRRR 336
QY 121 GDSRGSLLSPRPVSYLYKSGSGGPLLCPSGHAGVGFRAAUCTRGVAKAVDFPVVSMETTM 180
DB 337 GDSRGSLLSPRPVSYLYKSGSGGPLLCPSGHAGVGFRAAUCTRGVAKAVDFPVVSMETTM 396
QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240
DB 397 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 456
QY 241 AYNKSHAGIDPNRTGVRTITTCAPTITSTYKFLADGGCGGAYDIIICDECHSTDSTT 300
DB 457 AYNKSHAGIDPNRTGVRTITTCAPTITSTYKFLADGGCGGAYDIIICDECHSTDSTT 516
QY 301 ILGIGTVLDOAETAGARLVVLTATATPGSVTVPHNPNEEVALSSTGRIPIFYKAIPLETI 360
DB 517 ILGIGTVLDOAETAGARLVVLTATATPGSVTVPHNPNEEVALSSTGRIPIFYKAIPLETI 576
QY 361 KGRGHLIFCHSKKKDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 420
DB 577 KGRGHLIFCHSKKKDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 636
QY 421 GDPDSVDNCTCVTQTVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGMGIYRFTVPG 480
DB 637 GDPDSVDNCTCVTQTVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGMGIYRFTVPG 696
QY 481 ERPSGMFDSVLCBYDAGCAWYELTPAETSRLRAYLNTPLPGLPVCODHLEFWSVFTGL 540
DB 697 ERPSGMFDSVLCBYDAGCAWYELTPAETSRLRAYLNTPLPGLPVCODHLEFWSVFTGL 756
QY 541 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPSDOMWKCLIRLKEPTLHGPTPLL 600
DB 757 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPSDOMWKCLIRLKEPTLHGPTPLL 816
QY 601 YRLGAVQNEVTTHTPIKTYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSVVLVGR 660
DB 817 YRLGAVQNEVTTHTPIKTYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSVVLVGR 876
QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLIQTAKQEA 720
DB 877 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLIQTAKQEA 936
QY 721 AAPVVEKWTLEAFWAKHMMNFISGIOYLAGLSTLPGNPAIASLMFTAITSPLTQIH 780
DB 937 AAPVVEKWTLEAFWAKHMMNFISGIOYLAGLSTLPGNPAIASLMFTAITSPLTQIH 996
QY 781 TLLFNILGGWVAQAOLAPPASAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 840
DB 997 TLLFNILGGWVAQAOLAPPASAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 1056
QY 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900
DB 1057 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 1116
QY 901 GNHVSPTHYVDESAAAARVTOILSSLTITOLLKILHOWINEDCSTPCSGSWLRDYMWDIC 960

```

Db 1117 GNHVSPTHVVPESDAAARVTQILSSLTITOLLKRLHQMINDCSTPCSSWLRDYMWIC 1176  
QY 961 TVLTDKFWLQSKLLPRLPGVPPFFCQCGYKGVWGRGDMQITPCGQAQITGHVXGSMR 1020  
Db 1177 TVLTDKFWLQSKLLPRLPGVPPFFCQCGYKGVWGRGDMQITPCGQAQITGHVXGSMR 1236  
QY 1021 IVGPRTCSNTHGTPPINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHYVTGM 1080  
Db 1237 IVGPRTCSNTHGTPPINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHYVTGM 1296  
QY 1081 TTDNVKFCQVPAPEFFTEVDGVRHLRYAPACKPLLRBEVTLVGLNQVLVGSQLPCEPE 1140  
Db 1297 TTDNVKFCQVPAPEFFTEVDGVRHLRYAPACKPLLRBEVTLVGLNQVLVGSQLPCEPE 1356  
QY 1141 PDVAULTSMITDPSHITATKARRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD 1200  
Db 1357 PDVAULTSMITDPSHITATKARRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD 1416  
QY 1201 LIEANLWRQBMGNGNITRVESENKVILDSPEPLQAEDEREVSPAILRRSRKFPBRAM 1260  
Db 1417 LIEANLWRQBMGNGNITRVESENKVILDSPEPLQAEDEREVSPAILRRSRKFPBRAM 1476  
QY 1261 PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1477 PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1536  
QY 1321 ELATKTFGSSSESSAVDSGTATASPOPSDDGAGSDVESYSMPLEGEPPDLSDGSM 1380  
Db 1537 ELATKTFGSSSESSAVDSGTATASPOPSDDGAGSDVESYSMPLEGEPPDLSDGSM 1596  
QY 1381 STVSEEAESDVVCCSMSTWTGALITPCAARETKLPINALNSLRLHNLVYATTSRAS 1440  
Db 1597 STVSEEAESDVVCCSMSTWTGALITPCAARETKLPINALNSLRLHNLVYATTSRAS 1656  
QY 1441 LRQKVTDFRLQVLDHYRDVLKEMKAKASTVKALLSVEBEACKLTPPHSARSKFYGAK 1500  
Db 1657 LRQKVTDFRLQVLDHYRDVLKEMKAKASTVKALLSVEBEACKLTPPHSARSKFYGAK 1716  
QY 1501 DVRLNSSKANVHRSWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
Db 1717 DVRLNSSKANVHRSWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1776  
QY 1561 VRUCEKVALYDVSTLPOAVMGSSYGFQSPQGVVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
Db 1777 VRUCEKVALYDVSTLPOAVMGSSYGFQSPQGVVEFLVNAWKAKCPMGFAYDTRCFDS 1836  
QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKGQNGYRRCRASGVLIT 1680  
Db 1837 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKGQNGYRRCRASGVLIT 1896  
QY 1681 TSCGNTLTCVYLKAAACRAAKLODCTMLVCGDDLVVICESAGTOEDEASLRAFTAMTRY 1740  
Db 1997 TSCGNTLTCVYLKAAACRAAKLODCTMLVCGDDLVVICESAGTOEDEASLRAFTAMTRY 1956  
QY 1741 SAPPGDPKPEYDLELITSCSSNVSVADHASKGVYVLTDRPTTPLARAAMETARHTPVN 1800  
Db 1957 SAPPGDPKPEYDLELITSCSSNVSVADHASKGVYVLTDRPTTPLARAAMETARHTPVN 2016  
QY 1801 SWLGNIIWYAPTLWARMILMTHFFSIIILAQOLEKALDCQIYGACYSIEPLDLPQIILQRL 1860  
Db 2017 SWLGNIIWYAPTLWARMILMTHFFSIIILAQOLEKALDCQIYGACYSIEPLDLPQIILQRL 2076  
QY 1861 HGLSAFSLHSVSGEINRVASCLKLGVPPLRWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 2077 HGLSAFSLHSVSGEINRVASCLKLGVPPLRWRHARSVRARLLSQGGRAATCGKYL 2136  
QY 1921 NWAVRTKLTLPPIPAASQDLSSMEVAGSGDIIYHSLSRAPRPFWMCLLLSVGVGIY 1980  
Db 2137 NWAVRTKLTLPPIPAASQDLSSMEVAGSGDIIYHSLSRAPRPFWMCLLLSVGVGIY 2196  
QY 1981 LLPNR 1985  
Db 2197 LLPNR 2201

## RESULT 6

US-10-789-355-3  
; Sequence 3, Application US/10789355  
; Publication No. US20040180333A1  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/789,355  
; CURRENT FILING DATE: 2004-02-27  
; PRIOR APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 882  
; OTHER INFORMATION: Xaa is Lys or Arg  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1489  
; OTHER INFORMATION: Xaa is Leu  
US-10-789-355-3

Query Match 99.8%; Score 10442; DB 16; Length 2201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPITAYSQQTGRLGCIITSLTGRDRNOVEGEVQVYVSTATQSFATCNGVCVTVVHGA 60  
Db 217 LAPITAYSQQTGRLGCIITSLTGRDRNOVEGEVQVYVSTATQSFATCNGVCVTVVHGA 276  
QY 61 GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPCARSILTPCTCGSSDLYLVTTHADVIVRRR 120  
Db 277 GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPCARSILTPCTCGSSDLYLVTTHADVIVRRR 336  
QY 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
Db 337 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 396  
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFGG 240  
Db 397 RSPVFTDNSPPAVPQTFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFGG 456  
QY 241 AYMSKAHGDIPNIRGTVRTITTCGAPITYSYGKFLADGGCSGGAYDIIICDECHSTDTST 300  
Db 457 AYMSKAHGDIPNIRGTVRTITTCGAPITYSYGKFLADGGCSGGAYDIIICDECHSTDTST 516  
QY 301 ILGIGTVLDOAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPFGKAIPIETI 360  
Db 517 ILGIGTVLDOAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPFGKAIPIETI 576  
QY 361 KGRHILFCHSKKKKDELAALKSLGLNAVAYVYRGDLVSVIPTSGDVIVVATDALMTGFT 420  
Db 577 KGRHILFCHSKKKKDELAALKSLGLNAVAYVYRGDLVSVIPTSGDVIVVATDALMTGFT 636  
QY 421 GDFSDVIDCNTCTVQTVDFSLDPTFTIETTVPQDVAVSQRGRGRTGRGMGIYRFVTPG 480  
Db 637 GDFSDVIDCNTCTVQTVDFSLDPTFTIETTVPQDVAVSQRGRGRTGRGMGIYRFVTPG 696  
QY 481 ERPSGMFDDSVLCECYDAGCAWVELTPAETSRLRAYLNTPLGVCDHLEFWESVPTGL 540  
Db 697 ERPSGMFDDSVLCECYDAGCAWVELTPAETSRLRAYLNTPLGVCDHLEFWESVPTGL 756





Db 337 GDSRGLSPRPVSYLKGSGCLLCPGSHAVGIFRAAVCTRGVAKAYDFVPVSMETTM 396  
Qy 181 RSPVPTDNSSPAVDQTTQVQVHLHAPTQSGSKTKVPAAYAAQGYKVLVLPNSVAATLFGF 240  
Db 397 RSPVPTDNSSPAVDQTTQVQVHLHAPTQSGSKTKVPAAYAAQGYKVLVLPNSVAATLFGF 456  
Qy 241 AYMSKAHGDINIRGTGRTITGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 457 AYMSKAHGDINIRGTGRTITGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 516  
Qy 301 ILGIGTVLQDAETAGARLWLATATPPGSVTVPHNIEEVALSSSTGEIPFYKGAIPJETI 360  
Db 517 ILGIGTVLQDAETAGARLWLATATPPGSVTVPHNIEEVALSSSTGEIPFYKGAIPJETI 576  
Qy 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Db 577 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 636  
Qy 421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQAVRSORRGRTGGRMGIVRFVTPG 480  
Db 637 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQAVRSORRGRTGGRMGIVRFVTPG 696  
Qy 481 ERPSGMFSSVLCBDCYDAGCAWYBELTPAETSVRLRAYLNTFGLPVCQDHLFEFWSVFTGL 540  
Db 697 ERPSGMFSSVLCBDCYDAGCAWYBELTPAETSVRLRAYLNTFGLPVCQDHLFEFWSVFTGL 756  
Qy 541 THIDAHFSLQTKQAGDNFPYLVAYQATVCARQAAPPSPDQWKCLRLKPTLHGPTPL 600  
Db 757 THIDAHFSLQTKQAGDNFPYLVAYQATVCARQAAPPSPDQWKCLRLKPTLHGPTPL 816  
Qy 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTVTSLVGLVLAALAAAYCLTTGSSVVI 660  
Db 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTVTSLVGLVLAALAAAYCLTTGSSVVI 876  
Qy 661 IILSGKPAIIPDRVLVREFDEMEECASHLPYIBQGMQLAQFQKQKXGILQATKQAEA 720  
Db 877 IILSGKPAIIPDRVLVREFDEMEECASHLPYIBQGMQLAQFQKQKXGILQATKQAEA 936  
Qy 721 AAPVVEKSWRTLEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPLTQ 780  
Db 937 AAPVVEKSWRTLEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPLTQ 996  
Qy 781 TLLPNILGGWVAQALAPPSAASAFVAGIAGAAVSGTGLGKVLVDIILAGYAGAGALVA 840  
Db 997 TLLPNILGGWVAQALAPPSAASAFVAGIAGAAVSGTGLGKVLVDIILAGYAGAGALVA 1056  
Qy 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 900  
Db 1057 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 1116  
Qy 901 GNVHSPHYYPESDAAARVTQILSSLIITQLKRLHOWINEDCSTPCSGWLRDWDWIC 960  
Db 1117 GNVHSPHYYPESDAAARVTQILSSLIITQLKRLHOWINEDCSTPCSGWLRDWDWIC 1176  
Qy 961 TVLTDFKTLQSKLLPRLPGVPFPSCORGKYGWVRGDIIMQTTCCPQCAOITGHVKNCSMR 1020  
Db 1177 TVLTDFKTLQSKLLPRLPGVPFPSCORGKYGWVRGDIIMQTTCCPQCAOITGHVKNCSMR 1236  
Qy 1021 IVGERTCSNTWHGTFPINAVYTTGCTPSPAPNYSRALWRVAAEYVEVTVRGDFHYVTGM 1080  
Db 1237 IVGERTCSNTWHGTFPINAVYTTGCTPSPAPNYSRALWRVAAEYVEVTVRGDFHYVTGM 1296  
Qy 1081 TTDNVKCPQVPAPEFEFTEVDGVLHRYAPACKPLLEBVTFLVGLNQYLVGSLPCEPE 1140  
Db 1297 TTDNVKCPQVPAPEFEFTEVDGVLHRYAPACKPLLEBVTFLVGLNQYLVGSLPCEPE 1356  
Qy 1141 PDVAVLTSMLTDPDSHITAETAKRLARGSPSPSLASSASQLSAPLSKATCTTTRHDSPAD 1200  
Db 1357 PDVAVLTSMLTDPDSHITAETAKRLARGSPSPSLASSASQLSAPLSKATCTTTRHDSPAD 1416  
Qy 1201 LIEANLLWRQEMGGNITRVSEKNKVILDSFEPLQAEDEDEREVSVPAILRRSKFFPRAM 1260  
Db 1417 LIEANLLWRQEMGGNITRVSEKNKVILDSFEPLQAEDEDEREVSVPAILRRSKFFPRAM 1476

Qy 1261 PIWARPDPYNPPLLSWKPDDYVPPVHGCPLPPAKAPIPPPRRKRTVVLSESTVSSALA 1320  
Db 1477 PIWARPDPYNPPLLSWKPDDYVPPVHGCPLPPAKAPIPPPRRKRTVVLSESTVSSALA 1536  
Qy 1321 ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSSMPPLEGEPPGDDPLSDGSW 1380  
Db 1537 ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSSMPPLEGEPPGDDPLSDGSW 1596  
Qy 1381 STVSEERASEDVVCCSMSTYWTGALITPCAAETKLPINALSNSLLRHHNLVATTSRAS 1440  
Db 1597 STVSEERASEDVVCCSMSTYWTGALITPCAAETKLPINALSNSLLRHHNLVATTSRAS 1656  
Qy 1441 LRQKVTFRDRLQVLDHYRDVLEKMAKASTVKAKLSVEBACLTPPHSARSKEFGYAK 1500  
Db 1657 LRQKVTFRDRLQVLDHYRDVLEKMAKASTVKAKLSVEBACLTPPHSARSKEFGYAK 1716  
Qy 1501 DVRLSSKANVHRSVWKDLLEDTEPIDTIMAKNEVFCVQPEKGGKPKARLIVFPDLG 1560  
Db 1717 DVRLSSKANVHRSVWKDLLEDTEPIDTIMAKNEVFCVQPEKGGKPKARLIVFPDLG 1776  
Qy 1561 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVFEFLVNAWKAKCPCMGFAYDTRCFDS 1620  
Db 1777 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVFEFLVNAWKAKCPCMGFAYDTRCFDS 1836  
Qy 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
Db 1837 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1896  
Qy 1681 TSCGNTITCYLKAACRAAKLODCTMLVCGDDLVWICESAGTOEDRASIAFTEAMTRY 1740  
Db 1897 TSCGNTITCYLKAACRAAKLODCTMLVCGDDLVWICESAGTOEDRASIAFTEAMTRY 1956  
Qy 1741 SAPPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTTRDPTPLARAAMETARHTPEVN 1800  
Db 1957 SAPPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTTRDPTPLARAAMETARHTPEVN 2016  
Qy 1801 SWLGNIIIMYAPTLMARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
Db 2017 SWLGNIIIMYAPTLMARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLPQIIQRL 2076  
Qy 1861 HGLSAPLSHSYSGEINRVASCLRKLVPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 2077 HGLSAPLSHSYSGEINRVASCLRKLVPPLRVWRHARSVRARLLSQGGRAATCGKYL 2136  
Qy 1921 NWAVRTKLTPTPAASQDLSSWFAVAGYGGDIYHLSLRARPFWFWCMLLLSVGVGIY 1980  
Db 2137 NWAVRTKLTPTPAASQDLSSWFAVAGYGGDIYHLSLRARPFWFWCMLLLSVGVGIY 2196  
Qy 1981 LLPNR 1985  
Db 2197 LLPNR 2201

## RESULT 8

US-10-085-476-2  
; Sequence 2, Application US/10085476  
; Publication No. US20020164722A1  
; GENERAL INFORMATION:  
; APPLICANT: De Francesco, Raffaele  
; APPLICANT: Tomei, Licia  
; APPLICANT: Behrens, Sven-Brik  
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE  
; TITLE OF INVENTION: RNA-DEPENDENT RNA POLYMERASE AND TERMINAL NUCLEOTIDYL  
; TITLE OF INVENTION: TRANSFERASE ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)  
; FILE REFERENCE: IT0002PCA  
; CURRENT APPLICATION NUMBER: US/10/085,476  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: 08/952,981  
; PRIOR FILING DATE: 1998-03-23  
; PRIOR APPLICATION NUMBER: PCT/IT96/00106  
; PRIOR FILING DATE: 1996-05-24  
; PRIOR APPLICATION NUMBER: RM95A000343





QY 1741 SAPPDPPKPYDLELITSCSSNVSAHDASGKRVYLLTRDPTTPLARAWEHARHTPVN 1800  
DB 2766 SAPPDPPKPYDLELITSCSSNVSAHDASGKRVYLLTRDPTTPLARAWEHARHTPVN 2825  
QY 1801 SWLGNIMYAPTLMARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 1860  
DB 2826 SWLGNIMYAPTLMARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 2885  
QY 1861 HGLSAPSLHSYSGEINRVASCLURKLGVPPLRVHRARSVRARLLSQSGRAATCGKYL 1920  
DB 2886 HGLSAPSLHSYSGEINRVASCLURKLGVPPLRVHRARSVRARLLSQSGRAATCGKYL 2945  
QY 1921 NNAVRKLLKLTPIPAASQDLSSWVAGYSGGDIYHLSGRPRWFWMCLLLSVGVGY 1980  
DB 2946 NNAVRKLLKLTPIPAASQDLSSWVAGYSGGDIYHLSGRPRWFWMCLLLSVGVGIN 3005  
QY 1981 LLPNR 1985  
DB 3006 LLPNR 3010  
RESULT 10  
US-10-259-275-40  
; Sequence 40, Application US/10259275  
; Publication No. US20030125541A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemon, Stanley M.  
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE  
; FILE REFERENCE: 265.0007 0120  
; CURRENT APPLICATION NUMBER: US/10/259,275  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US 60/171,909  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: US 09/747,419  
; PRIOR FILING DATE: 2000-12-23  
; PRIOR APPLICATION NUMBER: US 60/325,236  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/338,123  
; PRIOR FILING DATE: 2001-11-13  
; NUMBER OF SEQ IDS: 73  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 40  
; LENGTH: 2985  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of SEQ ID  
; OTHER INFORMATION: NO:39  
US-10-259-275-40  
Query Match 92.8%; Score 9710; DB 14; Length 2985;  
Best Local Similarity 92.9%; Pred. No. 0;  
Matches 1831; Conservative 59; Mismatches 66; Indels 14; Gaps 3;  
QY 25 RDNQVEGEVQVSTATQSFATCVGCVTVVHGAGSKTLAGPKGPIQMYTNVDDLV 84  
DB 1021 RDNQVEGEVQVSTATQSFATCVGCVTVVHGAGSKTLAGPKGPIQMYTNVDDLV 1080  
QY 85 GWQAPPQARSUTPTCGSSDLYLVRHADVIPVRRRGRSGSLSPRVSYLKGSSGGL 144  
DB 1081 GWQAPPQARSUTPTCGSSDLYLVRHADVIPVRRRGRSGSLSPRVSYLKGSSGGL 1140  
QY 145 LCPSGHAGVGLFRAAVCTRGVAKAVDFVPVSMETMTMSPVFTDNSSPPAVPQTFQVAHLH 204  
DB 1141 LCPSGHAGVGLFRAAVCTRGVAKAVDFVPVSMETMTMSPVFTDNSSPPAVPQTFQVAHLH 1200  
QY 205 APTGSGKSTKVPAAAYAAQGYKVLNPSVAATLGFAGYMSKAHGIDNIRTVGRTITGA 264  
DB 1201 APTGSGKSTKVPAAAYAAQGYKVLNPSVAATLGFAGYMSKAHGIDNIRTVGRTITGA 1260  
QY 265 PITVSTYKFLADGGCGSGGAYDIIICDECHSTDTSTILGIGTWLDQAGTAGARLVVLATA 324

DB 1261 PITVSTYKFLADGGCGSGGAYDIIICDECHSTDTSTILGIGTWLDQAGTAGARLVVLATA 1320  
QY 325 TTPGGSVTVPHNPIEVALSSTGEIIPYKGAIPETIKGGRHLIFCHSKKCOELAKLSG 384  
DB 1321 TTPGGSVTVPHNPIEVALSSTGEIIPYKGAIPETIKGGRHLIFCHSKKCOELAKLSG 1380  
QY 385 LGINAVAYYRGDLDSVIPSIGDVIIVVATDALTGTGDFSDSVIDCMTCVTQTVDVSLDPT 444  
DB 1381 LGINAVAYYRGDLDSVIPSIGDVIIVVATDALTGTGDFSDSVIDCMTCVTQTVDVSLDPT 1440  
QY 445 FTIETTTPQDVAVSQRGRGTRGMGIYRFTVTPGRSPGMPDSSVLCRCYDAGCAWYE 504  
DB 1441 FTIETTTPQDVAVSQRGRGTRGMGIYRFTVTPGRSPGMPDSSVLCRCYDAGCAWYE 1500  
QY 505 LTPAETSIVLRAYLNTPLGVQDHLFWESVFTGLTHIDAHFLSOTKQAGDNFFYLVA 564  
DB 1501 LTPAETSIVLRAYLNTPLGVQDHLFWESVFTGLTHIDAHFLSOTKQAGDNFFYLVA 1560  
QY 565 QATVCARAOAPPPSDQMWKCLIRLXPTLHGPTPLLYRLGAVQNEVTTTTHTPIKYIMACM 624  
DB 1561 QATVCARAOAPPPSDQMWKCLIRLXPTLHGPTPLLYRLGAVQNEVTTTTHTPIKYIMACM 1620  
QY 625 SADLEVVTSTWLVGCVLAALAAAYCLTTGTSVVIIVGRIILSGKPAIIPDRVILVREDEME 684  
DB 1621 SADLEVVTSTWLVGCVLAALAAAYCLTTGTSVVIIVGRIILSGKPAIIPDRVILVREDEME 1680  
QY 685 ECASHLPYIEQGMOLTAEOFKQKALGLQATQKAEAAAPVVEKSKWTLAEAFNAKHMNFI 744  
DB 1681 ECASHLPYIEQGMOLTAEOFKQKALGLQATQKAEAAAPVVEKSKWTLAEAFNAKHMNFI 1740  
QY 745 SGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOHTLFLNIGWVAQAALPSSAASAF 804  
DB 1741 SGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOHTLFLNIGWVAQAALPSSAASAF 1800  
QY 805 VGAGIAGAAVSGISGLKVLVDILAGVAGVAGALVAFKVMGSGMPSTEDLVNLLPAILSP 864  
DB 1801 VGAGIAGAAVSGISGLKVLVDILAGVAGVAGALVAFKVMGSGMPSTEDLVNLLPAILSP 1860  
QY 865 GALVGVGCAAILRRHVGPGEVAVQMNRLIAFASRGNHVSPTHYVYPESDAAARVQTILS 924  
DB 1861 GALVGVGCAAILRRHVGPGEVAVQMNRLIAFASRGNHVSPTHYVYPESDAAARVQTILS 1920  
QY 925 SLTITQLLKLHOWINEDCSTPCSGSLRDVDMICTVLTDFKTLWLSKLLPRLPGVFF 984  
DB 1921 SLTITQLLKLHOWINEDCSTPCSGSLRDVDMICTVLTDFKTLWLSKLLPRLPGVFF 1980  
QY 985 SCORGKGVWRGDIWOTTCPCGAQITGHVKGSMRIIVGPRCTSNWHTGFFINAYTTGP 1044  
DB 1981 SCORGKGVWRGDIWOTTCPCGAQITGHVKGSMRIIVGPRCTSNWHTGFFINAYTTGP 2040  
QY 1045 CTSPAPNYSRALRWVAABRYEVTRVGDPHYVTGMTDNVKGPCQVPAPEFFTEVDGVR 1104  
DB 2041 CTSPAPNYSRALRWVAABRYEVTRVGDPHYVTGMTDNVKGPCQVPAPEFFTEVDGVR 2100  
QY 1105 LHRYPACKPLREEVTFVLGNLYVLSQ-----PCEPEPDVAVLTSMLTDPSSHITAE 1159  
DB 2101 LHRYPACKPLREEVTFVLGNLYVLSQ-----PCEPEPDVAVLTSMLTDPSSHITAE 2155  
QY 1160 TAKRRILARGSPSLASSASQLSAPSLKATCTTRH-----DSPADLIEANLLWROEMGN 1215  
DB 2156 LVGSQARGSPSLASSASQLSAPSLKATCTTRH-----DSPADLIEANLLWROEMGN 2215  
QY 1216 ITRVESNKKVILDSFEPLQAEEDEREVSVAETILRRSRKFFRAMPPIWARPDYNNPPLLES 1275  
DB 2216 ITRVESNKKVILDSFEPLQAEEDEREVSVAETILRRSRKFFRAMPPIWARPDYNNPPLLES 2275  
QY 1276 WKPDYVVPVPHGCPPLPAKAPPIPPRRKRTVVLSESTVSSALAEATKTFGSESASV 1335  
DB 2276 WKPDYVVPVPHGCPPLPAKAPPIPPRRKRTVVLSESTVSSALAEATKTFGSESASV 2335  
QY 1336 DSGTATASDPOPSDDGADGSDVSESSMPPLEGSPGPDLSGDSWSTVSEASDVVCCS 1395

Db 2336 DSGTATAPDQTSNKGKSDAESCMMPELGEPCDPLSDGSMSTVSEBAGESVVCSS 2395  
QY 1396 NSYTWGTGALITPCAAEETKLPINALSNSLLRHNLVYATTSASLRQKKVTFDRLOVLD 1455  
Db 2396 NSYTWGTGALITPCAAEESKLPINALSNSLLRHNLVYATTSASLRQKKVTFDRLOVLD 2455  
QY 1456 DHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYKADVRNLSSKAVNHHS 1515  
Db 2456 DHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYKADVRNLSSRAINHHS 2515  
QY 1516 VKOLLEDETETIDITINAKNEVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVVST 1575  
Db 2516 VVEDLLEDTVPIDITINAKNEVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVVST 2575  
QY 1576 LPOAVMGSSYGFQSPGQRFVFLYNAWAKKCPMGFAVDTRCFDSTVTENDIRVEESYQ 1635  
Db 2576 LPOAVMGSSYGFQSPGQRFVFLYNAWAKKCPMGFSDYTRCFDSTVTENDIRVEESYQ 2635  
QY 1636 CCDLAPEARQAIRSLTERLYTGGPLTNSKGQCYRRRCRAGSVLTSCGNTLTCYLKAAA 1695  
Db 2636 CCDLAPEARQAKSLTERLYTGGPLTNSKGQCYRRRCRAGSVLTSCGNTLTCYLKASA 2695  
QY 1696 ACRAAKIQCOTMLVCGDOLVWICESAGTQDEDEASLRAFTTEAMTRYSAAPPDPPKPEYDLE 1755  
Db 2696 ACRAAKIQCOTMLVCGDOLVWICESAGTQDEDAASLRVFTTEAMTRYSAAPPDLPQPEYDLE 2755  
QY 1756 LITSCSSNVSAHDASGKRVYLTDPPTPLARAAWETARHTPVNSWLNIIIMVAPTUWA 1815  
Db 2756 LITSCSSNVSAHDASGKRVYLTDPPTPLARAAWETARHTPVNSWLNIIIMVAPTUWA 2815  
QY 1816 RMILMTHFFSILLAQEKLKALDCQIYACYSIEPLDLPQIIORLHGLSAFSLHSYSPE 1875  
Db 2816 RMILMTHFFSILLAQEKLKALDCQIYACYSIEPLDLPQIIORLHGLSAFSLHSYSPE 2875  
QY 1876 INRVASCLKUGVPLRVWRHARSVRARLLSQGGRATCGKILFNMAVTKLITPIPA 1935  
Db 2876 INRVASCLKUGVPLRVWRHARSVRARLLSQGGRATCGKILFNMAVTKLITPIPA 2935  
QY 1936 ASQDLASWFWAGYSGGDIYHSLSRARPRWMLCLLLSVGVGIYLLPNNR 1985  
Db 2936 ASQDLASWFWAGYSGGDIYHSLSRARPRWMLCLLLSVGVGIYLLPNNR 2985

## RESULT 11

US-10-296-734-406  
; Sequence 406, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU P07761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 406  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: HepC la consensus polyprotein  
US-10-296-734-406

Query Match 90.3%; Score 9454.5; DB 15; Length 3011;  
Best Local Similarity 88.08; Pred. No. 0;  
Matches 1749; Conservative 134; Mismatches 101; Indels 3; Gaps 3;  
QY 1 MAPITAYSQTRGLGCIITSLTGRDNQVEGEVQVYVSTATQSFATCVNGVCWTVYHGA 60  
Db 1026 LAPITAYAQTRGLGCIITSLTGRDNQVEGEVQVYVSTATQSFATCVNGVCWTVYHGA 1085

QY 61 GSKTLAGKGPITOMYTNVDQLVGWQAPPGARSLTPTCTCGSSDLVLYTRHADVIPVRRR 120  
Db 1086 GTRTIASPKGPVIMYTNVDQLVGWQAPPGARSLTPTCTCGSSDLVLYTRHADVIPVRRR 1145  
QY 121 GDSRGSLLSPRPVSVLKSSGGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPEVSEMETM 180  
Db 1146 GDSRGSLLSPRPVSVLKSSGGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPEVSEMETM 1205  
QY 181 RSPVFTDNSSPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQYKVLVLPNSVAATLPG 240  
Db 1206 RSPVFTDNSSPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQYKVLVLPNSVAATLPG 1265  
QY 241 AYMSKAHGDINIRTVRTITGAPITYSTYKFLADGGCGSGAYDIIICDECHSTDTT 300  
Db 1266 AYMSKAHGDINIRTVRTITGSPITYSTYKFLADGGCGSGAYDIIICDECHSTDTT 1325  
QY 301 ILGIGTVLDQAEAGARLWILATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 360  
Db 1326 ILGIGTVLDQAEAGARLWILATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 1385  
QY 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIWVATDALMTGFT 420  
Db 1386 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSVIPTSGDVIWVATDALMTGFT 1445  
QY 421 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRRQRTGRGRMGIYRFVTPG 480  
Db 1446 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRRQRTGRGRMGIYRFVTPG 1505  
QY 481 ERPSGMDSSVLCBYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWSVFTGL 540  
Db 1506 ERPSGMDSSVLCBYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWSVFTGL 1565  
QY 541 THIDAHFLSOTKQAGDNPPYLVAYQATVCARAQAPPPSWDMWCKLIRKLPTLHGPTLL 600  
Db 1566 THIDAHFLSOTKQAGDNPPYLVAYQATVCARAQAPPPSWDMWCKLIRKLPTLHGPTLL 1625  
QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTTSTWLVGGVLAALAAAYCLTTGTVVIVGR 660  
Db 1626 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTTSTWLVGGVLAALAAAYCLTTGTVVIVGR 1685  
QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIISQGMOLABQPKQKALGLIQTATQAEA 720  
Db 1686 IILSGKPAIIPDREVLVREFDEMEECASHLPYIISQGMOLABQPKQKALGLIQTATQAEA 1745  
QY 721 AAPVVEKSWRTLEAFNAKHMNFISGQYLAGLSTLPCNPAIASLMAFTASITPLTTQH 780  
Db 1746 IAPAVQINWQKLEVFNAKHMNFISGQYLAGLSTLPCNPAIASLMAFTASITPLTTQH 1805  
QY 781 TLLFNILGGWVAQOLAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQOLAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1865  
QY 841 FKMSGEMSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEVAVWNRILIAFASR 900  
Db 1866 FKMSGEMSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEVAVWNRILIAFASR 1925  
QY 901 GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCSTPCSGWLRDWDWIC 960  
Db 1926 GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCSTPCSGWLRDWDWIC 1985  
QY 961 TVLTDFKWLQSKLLPRLPGVPPFFSCQRYKGVWRGDGIMQTTCPGCAQITGHVANGSMR 1020  
Db 1986 EVLSDFKWLKAKLIMPOLPGIPFVSCQRYKGVWRGDGIMHTRCHGCAEITGHVANGTMR 2045  
QY 1021 IVGPRICSNWHTGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTVRGDFHYVTGM 1080  
Db 2046 IVGPRICSNWHTGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTVRGDFHYVTGM 2105  
QY 1081 TTDNVCPCQVPAPEFTEVDGVRHLHRYAPACKPLLEEVTVFLVGLNQYLVGSOLPCEPE 1140  
Db 2106 TTDNVCPCQVPAPEFTEVDGVRHLHRYAPACKPLLEEVTVFLVGLNQYLVGSOLPCEPE 2165

QY	1141	PDVAVLTSM	LTDP	SHITAE	TAKR	RLARG	PPSI	ASS	SQ	LSAPS	KA	TKT	RHDS	SP	DAD	1200
Db	2166	PDVAVLTSM	LTDP	SHITAE	AAGR	RLARG	PPSP	MASS	AS	Q	LSAP	KA	TKT	AN	HDSP	DAE
QY	1201	LI	AN	LLWR	QEM	GN	IT	RV	SE	NK	VIL	D	SF	EL	QAE	DE
Db	2226	LI	AN	LLWR	QEM	GN	IT	RV	SE	NK	VIL	D	SF	EL	QAE	DE
QY	1261	PI	WA	R	PD	YN	P	LL	ES	WK	D	P	Y	P	V	P
Db	2286	P	WA	R	PD	YN	P	LL	ES	WK	D	P	Y	P	V	P
QY	1321	EL	AT	K	P	G	S	E	S	A	V	D	S	T	A	P
Db	2346	EL	AT	K	P	G	S	E	S	T	A	P	D	S	T	A
QY	1380	W	S	T	V	S	E	A	-	S	E	D	V	V	C	C
Db	2405	W	S	T	V	S	E	A	G	T	E	D	V	V	C	C
QY	1439	A	S	I	R	Q	K	V	T	F	R	I	Q	V	L	D
Db	2465	A	C	O	R	Q	K	V	T	F	R	I	Q	V	L	D
QY	1499	A	K	O	V	R	N	L	S	K	A	V	H	I	R	S
Db	2525	A	K	O	V	R	C	H	A	K	A	V	H	I	R	S
QY	1559	L	G	V	R	C	E	K	M	A	L	D	V	V	S	T
Db	2585	L	G	V	R	C	E	K	M	A	L	D	V	V	S	T
QY	1619	D	S	T	V	T	E	N	D	I	R	V	E	S	I	Y
Db	2645	D	S	T	V	T	E	S	D	I	R	V	E	S	I	Y
QY	1679	L	T	T	S	C	G	N	T	L	C	Y	L	A	A	A
Db	2705	L	T	T	S	C	G	N	T	L	C	Y	L	A	A</	

RESULT 12

RESULT 12  
IIS-09-742-659-4

US-09-742-659-4  
: Sequence 4 Application IIS/09742659

; Sequence 4, Application US, Patent No. US20010034018A1

; Patent No. US2001003

GENERAL INFORMATION:

APPLICANT: Hong, Zhi

APPLICANT: Butkiewicz, Na

;; APPLICANT: BUCKIEWICZ, NANCY  
;; APPLICANT: Zhong, Weidong

APPLICANT: Zhong, weidong  
APPLICANT: Ingravalle, Paul

APPLICANT: Wright - Minogue, Jacm

APPLICANT: Wright-Minogue, Jacques  
APPLICANT: Paul Johnson v

```

; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
;
; FILE REFERENCE: ID01116
;
; CURRENT APPLICATION NUMBER: US/09/742,659
;
; CURRENT FILING DATE: 2000-12-21
;
; PRIOR APPLICATION NUMBER: US 60/171,469
;
; PRIOR FILING DATE: 1999-12-22
;
; NUMBER OF SEQ ID NOS: 16
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
;
; LENGTH: 3011
;
; TYPE: PRT
;
; ORGANISM: Hepatitis C virus
;
; US-09-742-659-4

```

Query Match.	90.2%;	Score	9434.5;	DB	9;	Length	3011;																																																		
Best local Similarity	87.9%;	Pred.	No. 0;																																																						
Matches	1746;	Conservative	135;	Mismatches	103;	Indels	3;																																																		
Gaps	3;																																																								
Qy	1	M	A	P	I	T	A	S	Q	O	T	R	G	L	L	G	C	I	I	T	S	L	G	R	N	O	V	E	G	V	Q	V	V	S	T	A	T	Q	S	F	L	A	T	C	V	N	G	W	T	V	H	G	A	60			
Db	1026	L	A	P	I	T	A	S	Q	O	T	R	G	L	L	G	C	I	I	T	S	L	G	R	N	O	V	E	G	V	Q	V	I	S	T	A	T	Q	I	F	L	A	T	C	I	N	G	W	T	V	H	G	A	1085			
Qy	61	G	S	K	T	L	A	G	P	K	P	I	T	M	N	V	D	O	D	I	V	G	W	A	P	G	A	S	L	T	P	C	T	C	G	S	S	D	L	V	L	T	H	A	D	V	I	P	V	R	R	120					
Db	1086	G	T	R	T	A	S	P	K	P	V	I	T	M	N	V	D	O	D	I	V	G	W	P	A	P	O	G	S	R	S	L	T	P	C	T	C	G	S	S	D	L	V	L	T	H	A	D	V	I	P	V	R	R	1145		
Qy	121	G	D	S	R	G	S	L	L	S	R	P	P	S	Y	L	K	G	S	G	G	P	L	C	P	S	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	D	V	P	V	P	V	E	S	M	E	T	T	180
Db	1146	G	D	S	R	G	S	L	L	S	R	P	P	S	Y	L	K	G	S	G	G	P	L	C	P	A	G	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	D	V	I	P	V	E	N	E	T	T	1205		
Qy	181	R	S	P	V	F	T	D	N	S	P	P	A	V	P	O	T	F	Q	V	A	H	L	H	A	P	T	G	S	K	S	T	K	V	P	A	A	Y	A	O	Q	K	V	L	V	L	N	P	S	V	A	A	T	L	G	F	240
Db	1206	R	S	P	V	F	T	D	N	S	P	P	A	V	P	O	S	Q	F	V	A	H	L	H	A	P	T	G	S	K	S	T	K	V	P	A	A	Y	A	O	Q	K	V	L	V	L	N	P	S	V	A	A	T	L	G	F	1265
Qy	241	A	Y	M	S	K	A	G	I	D	N	I	R	T	G	V	R	T	I	T	G	A	P	I	T	S	Y	T	G	K	F	L	A	D	G	G	S	G	G	A	Y	D	I	I	C	D	E	C	H	S	T	D	T	300			
Db	1266	A	Y	M	S	K	A	G	V	D	N	I	R	T	G	V	R	T	I	T	G	S	P	I	T	S	Y	T	G	K	F	L	A	D	G	G	S	G	G	A	Y	D	I	I	C	D	E	C	H	S	T	D	T	1325			
Qy	301	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	P	G	S	V	T	V	P	H	P	N	E	E	V	A	L	S	T	G	E	I	P	F	Y	G	K	A	I	P	L	E	T	360		
Db	1326	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	P	G	S	V	T	V	P	H	P	N	E	E	V	A	L	S	T	G	E	I	P	F	Y	G	K	A	I	P	L	E	V	I	1385	
Qy	361	K	G	R	H	L	I	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	N	A	V	A	Y	R	G	L	D	V	S	I	P	S	G	D	V	I	V	A	T	D	A	L	M	T	G	F	420					
Db	1386</																																																								

QY 841 FKVMGEMPESTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 900  
Db 1866 FKIMGEVPESTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 1925  
QY 901 GNHVSPTHVVPESDAAARVTQLLSITITQLLKRHLHOWINEDCSTPCSGSWLRDWDWIC 960  
Db 1926 GNHVSPTHVVPESDAAARVTQLLSITITQLLKRHLHOWINEDCSTPCSGSWLRDWDWIC 1985  
QY 961 TVLTDFKWLQSKLLPRLPDGVFFSCQRYGKQVWRGDGIMQITPCGAQITGHVKNRMSR 1020  
Db 1986 EYLSDFKTLWKAKLPQLPFVSCQRYGKQVWRGDGIMHTRCHCGAIBTGHVKNRMSR 2045  
QY 1021 IVGPRTCSNTHGTTPINAYTTGCTPSPAPNYSALMRVAAEEVEVTRVGDHFHYTGM 1080  
Db 2046 IVGPRTCSNTHGTTPINAYTTGCTPSPAPNYSALMRVAAEEVEVTRVGDHFHYTGM 2105  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRHLRYAPACKPLREEVTLVLGNOLVGSQPCPE 1140  
Db 2106 TTDNVKCPQVPAPEFFTEVDGVRHLRYAPACKPLREEVTLVLGNOLVGSQPCPE 2165  
QY 1141 PDVAVLTSMIDPSHITAEAKRLARGSPSSASSASQSLKATCTTTRHSDPAD 1200  
Db 2166 PDVAVLTSMIDPSHITAEAKRLARGSPSSASSASQSLKATCTTTRHSDPAD 2225  
QY 1201 LIEANLLRWQMGNI TRVESNKVVIDSPEPLQAEEDEREVSYPAILRRSRKPPRAM 1260  
Db 2226 LIEANLLRWQMGNI TRVESNKVVIDSPEPLQAEEDEREVSYPAILRRSRKPPRAM 2285  
QY 1261 PIWAPDYNPPLLESWKDPDYPVPHGCPAPKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 2286 PIWAPDYNPPLLESWKDPDYPVPHGCPAPKAPPIPPRRKRTVVLSESTVSSALA 2345  
QY 1321 ELATKTFGSSSSAVDSGTATSPQPSDDG-DAGSDVESYSSMPLEGECDPLSDGS 1379  
Db 2346 ELATKTFGSSSSAVDSGTATSPQPSDDG-DAGSDVESYSSMPLEGECDPLSDGS 2404  
QY 1380 WSTVSEEA-SEDVVCSSSYMTGALITPCAETKLPINALNSLLRHHNLVATTSRS 1438  
Db 2405 WSTVSSGADTEDVVCSSSYMTGALITPCAETKLPINALNSLLRHHNLVATTSRS 2464  
QY 1439 ASLRQKKVTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLPHPHSARSKFYG 1498  
Db 2465 ACQKQKVTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLPHPHSARSKFYG 2524  
QY 1499 AKOVNLTSSKAVNHRSVWKDLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 1558  
Db 2525 AKOVNLTSSKAVNHRSVWKDLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 2584  
QY 1559 LGVRVCEKMALDYVSTLPQAVMGSSYGFQYSGORVEFLYNWAKKCPMPAYDTRCF 1618  
Db 2585 LGVRVCEKMALDYVSTLPQAVMGSSYGFQYSGORVEFLYNWAKKCPMPAYDTRCF 2644  
QY 1619 DSTVTENDIRVEESIYQCCDLAPARQAIRSLTBLTYIGGPLTNSKGNCCGYRRCRASGV 1678  
Db 2645 DSTVTENDIRVEESIYQCCDLAPARQAIRSLTBLTYIGGPLTNSKGNCCGYRRCRASGV 2704  
QY 1679 LTTCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEBASLRAFTAMT 1738  
Db 2705 LTTCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEBASLRAFTAMT 2764  
QY 1739 RYSAPPGDPKPEYDLELITSCSSNVSAHNDASGRVYVLTREDPTPLARAAMTARHTP 1798  
Db 2765 RYSAPPGDPKPEYDLELITSCSSNVSAHNDASGRVYVLTREDPTPLARAAMTARHTP 2824  
QY 1799 VNSWLGNIIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGCYSIEPLDPLQIQ 1858  
Db 2825 VNSWLGNIIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGCYSIEPLDPLQIQ 2884  
QY 1859 RLHGLSFLSHSYSGEINRVASCLRLGVPLPLWRHRRARSVEARLLSQGGRAATCGKY 1918  
Db 2885 RLHGLSFLSHSYSGEINRVASCLRLGVPLPLWRHRRARSVEARLLSQGGRAATCGKY 2944

QY 1919 LFNWAVRTKLTPTIPAAASQDLSSWFWAGYSGGDIYHSLSRARPRFWMCILLISVGVG 1978  
Db 2945 LFNWAVRTKLTPTIPAAASQDLSSWFWAGYSGGDIYHSLSRARPRFWMCILLISVGVG 3004  
QY 1979 IYLLPNR 1985  
Db 3005 IYLLPNR 3011  
RESULT 13  
US-09-891-894-3  
; Sequence 3, Application US/09891894  
; Publication No. US20030013081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, William  
; APPLICANT: Maddon, Paul  
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INFI  
; FILE REFERENCE: 2048/64896/JPW/SHS  
; CURRENT APPLICATION NUMBER: US/09/891,894  
; CURRENT FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: hepatitis c virus  
US-09-891-894-3

Query Match 90.1%; Score 9430.5; DB 10; Length 3011;  
Best Local Similarity 87.8%; Pred. No. 0;  
Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;  
QY 1 MAPITASYQQRTGLLIGCIITSLTCGRDRNQVEGEVQVWSTATQSLATCVNGVCWTVYHGA 60  
Db 1026 LAPITAYAAQRTGLLIGCIITSLTCGRDRNQVEGEVQVWSTATQSLATCVNGVCWTVYHGA 1085  
QY 61 GSKTLAGPKBITOMYTNVDQDLVGMQAPPGARSLLTPCTCGSSDLYLVTRHADVIPVRRR 120  
Db 1086 GTRTIASPKGPIVOMYTNVDQDLVGMQAPPGARSLLTPCTCGSSDLYLVTRHADVIPVRRR 1145  
QY 121 GDSRGSLLSPRPVSVLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 1146 GDSRGSLLSPRPVSVLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
QY 181 RSPVFTDNSSPPVAPQFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVNPVSAATLGF 240  
Db 1206 RSPVFTDNSSPPVAPQFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVNPVSAATLGF 1265  
QY 241 AYMSKAHGDINIRTVRTITGAPITVSTYKFLADGGCGGAYDIIICDECHSTDST 300  
Db 1266 AYMSKAHGDINIRTVRTITGAPITVSTYKFLADGGCGGAYDIIICDECHSTDST 1325  
QY 301 ILGIGTVLDDQAEATAGARLVLATATPGSVTVPHNIEEVALSSTGEIPFGKAIPIETI 360  
Db 1326 ILGIGTVLDDQAEATAGARLVLATATPGSVTVPHNIEEVALSSTGEIPFGKAIPIETI 1385  
QY 361 KGRHLLIFCHSKKKDELAALAKSLGLNVAAYRGLDVSVIPTSGDVIWVATDALMTGFT 420  
Db 1386 KGRHLLIFCHSKKKDELAALAKSLGLNVAAYRGLDVSVIPTSGDVIWVATDALMTGFT 1445  
QY 421 GDFSDVIDCNTCTVQTVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRGMGIYRFVTPG 480  
Db 1446 GDFSDVIDCNTCTVQTVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRGMGIYRFVTPG 1505  
QY 481 ERSGMFDSSLCECYDAGCAWYELPAETSVPLRAVINTPLGVCDHLEFESVETGL 540  
Db 1506 ERSGMFDSSLCECYDAGCAWYELPAETSVPLRAVINTPLGVCDHLEFESVETGL 1565  
QY 541 THIDAHFLSQTQKAGNFPYLVAYQATVCARAQAPPPSWDMKCLIRLKTPLHGPTELL 600  
Db 1566 THIDAHFLSQTQKAGNFPYLVAYQATVCARAQAPPPSWDMKCLIRLKTPLHGPTELL 1625  
QY 601 YRLGAVQNEVTTHTPIITKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLITGTSVVIVGR 660



Db 1626 YRLGAVQNEVTLTHPTIKITIMTKMSADLEWVTSWLVGGVLAALAAAYCLSTCCVIVGR 1685  
 QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQALAEQKQKAIGLLOPATKQAEA 720  
 Db 1686 IIVLSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQALAEQKQKAIGLLOPATKQAEA 1745  
 QY 721 AAPVSEKRTLEAFWAKHWNFIISGLOIAGLSTLPGNPAIASIAFAFTASITSPITTOH 780  
 Db 1746 ITPAVQTNWQKLEVFWAKHWNFIISGLOIAGLSTLPGNPAIASIAFAFTASITSPITTOH 1805  
 QY 781 TLLFNILGHWAAQALAPPASAFVAGIAGAAGVSGIGLKVLDIILAGYAGAGVAGALVA 840  
 Db 1806 TLLFNILGHWAAQALAPPASAFVAGIAGAAGVSGIGLKVLDIILAGYAGAGVAGALVA 1865  
 QY 841 FKYMSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAGVQMMNRLIAPASR 900  
 Db 1866 FKYMSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAGVQMMNRLIAPASR 1925  
 QY 901 GNVSPHYVPESDAARVQILSSLTITOLKRLHOWINEDCSTPCSGSLWLDVMDWIC 960  
 Db 1926 GNVSPHYVPESDAARVQILSSLTITOLKRLHOWINEDCSTPCSGSLWLDVMDWIC 1985  
 QY 961 TVLTDFTKWLQSKLPLPGVPFPFCQYKGVWKGDMQTTCPGCAQITGHVYKSGMR 1020  
 Db 1986 EVLSDFKTLKWLKXLMQPLGIPFVSCQYGVWKGDMQTTCPGCAQITGHVYKSGMR 2045  
 QY 1021 IVERTCSTNHWGTFPPINAVTTCGTPSPAPNYSRALWRVAABEYVETRVGFHYVTGM 1080  
 Db 2046 IVERTCSTNHWGTFPPINAVTTCGTPSPAPNYSRALWRVAABEYVETRVGFHYVTGM 2105  
 QY 1081 TTDNVKCPQVPAPEPTEVDGLVHLYAPACKPLLEETVGLNQLVGLSOLPCEPE 1140  
 Db 2106 TTDNVKCPQVPAPEPTEVDGLVHLYAPACKPLLEETVGLNQLVGLSOLPCEPE 2165  
 QY 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSPSSASSASQSLSPKATCTTTRHSDPAD 1200  
 Db 2166 PDVAVLTSMLTDSHITAEAKRLARGSPSPSSASSASQSLSPKATCTTTRHSDPAD 2225  
 QY 1201 LIBANLWROEMGNITRVSEKKNVILDSPEIQAEDEREVSVPABEILRRSKFRFAM 1260  
 Db 2226 LIBANLWROEMGNITRVSEKKNVILDSPEIQAEDEREVSVPABEILRRSKFRFAM 2285  
 QY 1261 PIWAPDYNPLLESKWDPDYPPVHVHGCPLPPAKAPPIDPPRRKRTVILSESIVSSALA 1320  
 Db 2286 PIWAPDYNPLLESKWDPDYPPVHVHGCPLPPAKAPPIDPPRRKRTVILSESIVSSALA 2345  
 QY 1321 ELATKTFGSSSAVDSTASPDQPSDDG-DAGSDVESYSSMPLEGEFGDPDLSDGS 1379  
 Db 2346 ELATKTFGSSSAVDSTASPDQPSDDG-DAGSDVESYSSMPLEGEFGDPDLSDGS 2404  
 QY 1380 WSTVSEEA-SEDDVCCSMSTWTGALLITPCAAETKLPINALSNLLRHNLVYATTSRS 1438  
 Db 2405 WSTVSSGADTEDVVCSSMSYWTGALVTPCAAEEQKLPINALSNLLRHNLVYATTSRS 2464  
 QY 1439 ASLRQKVTFTDLQVLDHVDVLEKEMKAKASTVKALLSVEEACKLTPPHSARSKTGYG 1498  
 Db 2465 ACQOKKVTFTDLQVLDHVDVLEKEMKAKASTVKALLSVEEACKLTPPHSARSKTGYG 2524  
 QY 1499 AKQVRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 1558  
 Db 2525 AKQVRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 2584  
 QY 1559 LGVRVCEKMAIYDVSTLPOAVSSYGFQYSPQRFVFLVNAWKAKCPMGFAVDTKCF 1618  
 Db 2585 LGVRVCEKMAIYDVSTLPOAVSSYGFQYSPQRFVFLVNAWKAKCPMGFAVDTKCF 2644  
 QY 1619 DSTVTENDIRVEESYOCCLDAPARQAIRLTERLYIGGPLTNSKONCGYRRCRASGV 1678  
 Db 2645 DSTVTENDIRVEESYOCCLDAPARQAIRLTERLYIGGPLTNSKONCGYRRCRASGV 2704  
 QY 1679 IITSCGNLTITCYLKAAAACRAAKIQDCTMLVCGDDLVVICSSAGTQDEASIRAFTEAMT 1738

Db 2705 LITSCGNLTITCYLKAAAACRAAKIQDCTMLVCGDDLVVICSSAGTQDEASIRAFTEAMT 2764  
 QY 1739 RYSAPPDGPDPKPEYDLELITSCSSNVSVADHAGKRVYLTROPTTPLAPAAWETARHTP 1798  
 Db 2765 RYSAPPDGPDPKPEYDLELITSCSSNVSVADHAGKRVYLTROPTTPLAPAAWETARHTP 2824  
 QY 1799 VNSWLGNIIMYAPTILWARMILMTHFFSILLAOEQLEKALDCQIYGACYSIEPDLPLQIQ 1858  
 Db 2825 VNSWLGNIIMYAPTILWARMILMTHFFSILLAOEQLEKALDCQIYGACYSIEPDLPLQIQ 2884  
 QY 1859 RLHGLSAFSLHSPGEINRVASCLKGLVPPPLRVWRHARSVRARLLSOGGAAATCGKY 1918  
 Db 2885 RLHGLSAFSLHSPGEINRVASCLKGLVPPPLRVWRHARSVRARLLSOGGAAATCGKY 2944  
 QY 1919 LFNWAVRTKLTPIPAASOLDLSSWVFVAGYSGDILVHLSRARPRFWMFMCLLLSVGVG 1978  
 Db 2945 LFNWAVRTKLTPIPAASOLDLSSWVFVAGYSGDILVHLSRARPRFWMFMCLLLSVGVG 3004  
 QY 1979 IYLLPNR 1985  
 Db 3005 IYLLPNR 3011

RESULT 14  
 US-10-184-150-3  
 : Sequence 3, Application US/10184150  
 : Publication No. US20030134297A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Olson, William  
 : APPLICANT: Maddon, Paul  
 : TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS IN  
 : FILE REFERENCE: 2048/64896-A/JFW/MAF/DJK  
 : CURRENT APPLICATION NUMBER: US/10/184,150  
 : CURRENT FILING DATE: 2002-12-10  
 : PRIOR APPLICATION NUMBER: 09/891,894  
 : PRIOR FILING DATE: 2001-06-26  
 : NUMBER OF SEQ ID NOS: 3  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 3  
 : LENGTH: 3011  
 : TYPE: PRT  
 : ORGANISM: Hepatitis C virus  
 US-10-184-150-3

Query Match 90.1%; Score 9430.5; DB 14; Length 3011;  
 Best Local Similarity 87.8%; Pred. No. 0;  
 Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;

Qy 1 MAPITAYSQOTRGLLGCIITSLTGRDRNQVEGEVQVYVSTATQSPATCNGVCTVTHGA 60  
 Db 1026 LAPITAYAQOTRGLLGCIITSLTGRDRNQVEGEVQVYVSTATQSPATCNGVCTVTHGA 1085  
 Qy 61 GSKTLAPKGPITOMYTNVDQDLVGVQAPPGARSLTCTCGSSDLVLTSHADVIPVRR 120  
 Db 1086 GTRTIASPKGPITOMYTNVDQDLVGVQAPPGARSLTCTCGSSDLVLTSHADVIPVRR 1145  
 Qy 121 GDSRGSLLSPRYSYLYKSGSGGGLLCPGSHAVGFRAAVCTRGVAKAVDFVFPVSEMTTM 180  
 Db 1146 GDSRGSLLSPRYSYLYKSGSGGGLLCPGSHAVGFRAAVCTRGVAKAVDFVFPVSEMTTM 1205  
 Qy 181 RSPVFTDNSPPAVPQTFQVAHLHAPITGSKSTKVPAAAYAAQYKVLVLPNSVAATLFGF 240  
 Db 1206 RSPVFTDNSPPAVPQTFQVAHLHAPITGSKSTKVPAAAYAAQYKVLVLPNSVAATLFGF 1265  
 Qy 241 AYNKSHAGIDPNIRTVRTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDTT 300  
 Db 1266 AYNKSHAGIDPNIRTVRTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDTT 1325  
 Qy 301 ILGIGTVLDQAEATAGARLVVLTATPPGSVTVPHNIEEVALSSTGEIPFYKGAIPETI 360  
 Db 1326 ILGIGTVLDQAEATAGARLVVLTATPPGSVTVPHNIEEVALSSTGEIPFYKGAIPETI 1385  
 Qy 361 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALTMTGFT 420

1386 KGRHLIFCHSKKCDLAAKVALGINAVAYRGLDVSIPTSGDVVVVSDALMTGFT 1445  
421 GDFSDVIDCNTCVTQTVFSLDPTTFTTTPQDAVSQRGRGTRGRMGYRFTVTPG 480  
1446 GDFSDVIDCNTCVTQTVFSLDPTTFTTTPQDAVSQRGRGTRGRMGYRFTVTPG 1505  
481 BRPSGMFSSVLCCEYDAGCAWELTPAETSVRLRAYLNTPLPYCODHLEFWESFTGL 540  
1506 BRPSGMFSSVLCCEYDAGCAWELTPAETSVRLRAYLNTPLPYCODHLEFWESFTGL 1565  
541 THIDAHFLSQTQAGDNFPYLVAQVATCARAQAPPPSDQWKKLIRLKPFLHGTPL 600  
1566 THIDAHFLSQTQAGDNFPYLVAQVATCARAQAPPPSDQWKKLIRLKPFLHGTPL 1625  
601 YRLGAVQNEVTTTHITIKYIMACMGADLEVVTSVWLVGVLAALAAAYCLTTGTSVVIVGR 660  
1626 YRLGAVQNEVTTTHITIKYIMACMGADLEVVTSVWLVGVLAALAAAYCLTTGTSVVIVGR 1685  
661 IILSGKPAIIPDREVLVREFDEMEBCASHLPYIEQGMOLAEQFKATGLLOTAKQEA 720  
1686 IILSGKPAIIPDREVLVREFDEMEBCASHLPYIEQGMOLAEQFKATGLLOTAKQEA 1745  
721 AAPVYESKRWLEAFWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSLTQH 780  
1746 ITPAVOTNWXKLEVPWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSLTQH 1805  
781 TLLFNILGWAQAOLAPSAASAFVAGIAGAAVSGIGIKVLVDILAGYGAGVAGALVA 840  
1806 TLLFNILGWAQAOLAPSAASAFVAGIAGAAVSGIGIKVLVDILAGYGAGVAGALVA 1865  
841 FKMSGEMPESTEDLVNLLPAILSPALVGVVCAAILRRHVGPBGAVQWNNRLTAFASR 900  
1866 FKMSGEMPESTEDLVNLLPAILSPALVGVVCAAILRRHVGPBGAVQWNNRLTAFASR 1925  
901 GNHVSPTHYVPESDAAARVTOILSLITLQILKRLHOMINEDCSTPCSGSWLRDWDWIC 960  
1926 GNHVSPTHYVPESDAAARVTOILSLITLQILKRLHOMINEDCSTPCSGSWLRDWDWIC 1985  
961 TVLTDFTKWLKSLPRLPVPFFSCQRYGVVWGDGIMQTTCPGQAQITGHVKNVSGMR 1020  
1986 EVLSDFKTLWAKLMPQFPIPVFVSCQRYGVVWGDGIMQTTCPGQAQITGHVKNVSGMR 2045  
1021 IVGPRTCSNTHGTFPFIINAYTGTCTPSPALVGVVCAAILRRHVGPBGAVQWNNRLTAFASR 1080  
2046 IVGPRTCSNTHGTFPFIINAYTGTCTPSPALVGVVCAAILRRHVGPBGAVQWNNRLTAFASR 2105  
1081 TTDNVKCPQVPAPEFFTEVDVGRVLRHYPACKPILLREEVFLVGLNOLVGLVGLPCPE 1140  
2106 TTDNVKCPQVPAPEFFTEVDVGRVLRHYPACKPILLREEVFLVGLNOLVGLVGLPCPE 2165  
1141 PDVAVLTSMLTDPHSHTAETAKRLARGSPSLASSASQSLASPLKATCTTRHSDPAD 1200  
2166 PDVAVLTSMLTDPHSHTAETAKRLARGSPSLASSASQSLASPLKATCTTRHSDPAD 2225  
1201 LIEANLLWRQMGNIIRVSENVKVVILDSPELQAEDEDEVSPAEILRRSRKFFRAM 1260  
2226 LIEANLLWRQMGNIIRVSENVKVVILDSPELQAEDEDEVSPAEILRRSRKFFRAM 1285  
1261 PIWARPDPNPLLESWKDPDYPVPHVGCPLPPAKAPPIDPPRRKRTVWISSEVSSALA 1320  
2286 PIWARPDPNPLLESWKDPDYPVPHVGCPLPPAKAPPIDPPRRKRTVWISSEVSSALA 2345  
1321 ELATKTCGSSSSAVDSGTATSPDQSDGDDG-DAGSVDSEYSSMPPIEGEPDPLSDGS 1379  
2346 ELATKTCGSSSSAVDSGTATSPDQSDGDDG-DAGSVDSEYSSMPPIEGEPDPLSDGS 2404  
1380 WSTVSEEA-SDVVCCSMYSWTGALITPCAABETKLPINALSNLLRHNLVYATTSRS 1438  
2405 WSTVSSGADTEDVVCCSMYSWTGALITPCAABETKLPINALSNLLRHNLVYATTSRS 2464  
1439 ASLFOKKVTFDLRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEGYG 1498

Db 2465 ACQROKKVTFDLRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEGYG 2524  
QY 1499 AKDVNLSKAVNHHSIRSVWVKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 1558  
Db 2525 AKDVNLSKAVNHHSIRSVWVKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 2584  
QY 1559 LGVRVCEKALYDVVSTLTPQAVMGSSYGFQSPGQVRVFLVNAWKAKKCPMGFAVDTRCF 1618  
Db 2585 LGVRVCEKALYDVVSTLTPQAVMGSSYGFQSPGQVRVFLVNAWKAKKCPMGFAVDTRCF 2644  
QY 1619 DSTVTENDIRVEESTIYQCDDLAPEARQAIIRSLTERLYTGGPLTNSKQNCQYRCRAGSV 1678  
Db 2645 DSTVTENDIRVEESTIYQCDDLAPEARQAIIRSLTERLYTGGPLTNSKQNCQYRCRAGSV 2704  
QY 1679 LTTSCGNTLTCYKAAAAACRAAKLQDCTMLVCGDDLVIWVICSAGTQDEEASLRAFEAMT 1738  
Db 2705 LTTSCGNTLTCYKAAAAACRAAKLQDCTMLVCGDDLVIWVICSAGTQDEEASLRAFEAMT 2764  
QY 1739 RYSAPPGPPPEYDLELITSCSSNVVAHDAGSKRVYLLTRDPTPLARAAMETARHTP 1798  
Db 2765 RYSAPPGPPPEYDLELITSCSSNVVAHDAGSKRVYLLTRDPTPLARAAMETARHTP 2824  
QY 1799 VNSMLGNIMVAPTLLWARMILMTHFFSILLAQEQLEKALDCOYGCACYSIEPLDLPQIIQ 1858  
Db 2825 VNSMLGNIMVAPTLLWARMILMTHFFSILLAQEQLEKALDCOYGCACYSIEPLDLPQIIQ 2884  
QY 1859 RLHGLSAFSLHSYSGPEINRVASCLRLGVPLRVHRARSVRARLLSOGGRAATCGKY 1918  
Db 2885 RLHGLSAFSLHSYSGPEINRVASCLRLGVPLRVHRARSVRARLLSOGGRAATCGKY 2944  
QY 1919 LFNWAVRTKLTLPIPAASQDLDSWVAGVSGGDIYHLSRARPRWFMWCLLLSVGVG 1978  
Db 2945 LFNWAVRTKLTLPIPAASQDLDSWVAGVSGGDIYHLSRARPRWFMWCLLLSVGVG 3004  
QY 1979 IYLLPNR 1985  
Db 3005 IYLLPNR 3011

RESULT 15  
US-10-328-997-3  
; Sequence 3, Application US/10328997  
; Publication No. US20030232745A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, William  
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INFI  
; FILE REFERENCE: 2048/54896-B/JPW/MAP/DJK  
; CURRENT APPLICATION NUMBER: US/10/328,997  
; CURRENT FILING DATE: 2002-12-24  
; PRIOR APPLICATION NUMBER: 09/891,894  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 10/184,150  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-10-328-997-3

Query Match 90.1%; Score 9430.5; DB 14; Length 3011;  
Best Local Similarity 87.8%; Pred. No. 0;  
Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;  
QY 1 MAPITAYSOQTRGLGCIITSLTGRDRNVEGEVQVSVSTATQSFATCVNGVCMVTVHGA 60  
Db 1026 LAPITAYAOQTRGLGCIITSLTGRDRNVEGEVQVSVSTATQSFATCVNGVCMVTVHGA 1085  
QY 61 GSKTLAPKGPITQMTYNTVDQDLVGVCAAPPGARSLLTCTCGSSDLVLTTHADVIPVRR 120  
Db 1086 GTRTIASPKGPVQMTYNTVDQDLVGVCAAPPGARSLLTCTCGSSDLVLTTHADVIPVRR 1145

QY	121	GBSRGSLSPRPVSYLKGSGGGLLCPSGHAGVIFRAAVCTRGVAKAVDPVPVSEMETTM	180
Db	1146	GBSRGSLSPRPVSYLKGSGGGLLCPAGHAVGLFRAAVCTRGVAKAVDPVFPVENLETTM	1205
QY	181	RSVPFTDNGSPPAVPOTFQVAHLHAPTGSKSTKVAAYAAQYKVLVLPNSVAATLFGF	240
Db	1206	RSVPFTDNGSPPAVPOSFQVAHLHAPTGSKSTKVAAYAAQYKVLVLPNSVAATLFGF	1265
QY	241	AYMSKAHGIDPNIRTVGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT	300
Db	1266	AYMSKAHGVDPNIRTVGRTITTCSPITYSTYKFLADGGCGGAYDIIICDECHSDATS	1325
QY	301	ILGIGVLDQAEATAGARLVVLAATATPPGSVTVPHPNIEVALSTGEIPYKGAIPLETI	360
Db	1326	ILGIGVLDQAEATAGARLVVLAATATPPGSVTVPHPNIEVALSTGEIPYKGAIPLEVI	1385
QY	361	KGRHILIFCHSKKCDLAAKLSGLNVAAYVRGLDVSIVPTSGDVIVVATDALMTGFT	420
Db	1386	KGRHILIFCHSKKCDLAAKLVAGLNVAAYVRGLDVSIVPTSGDVIVVATDALMTGFT	1445
QY	421	GFDSVIDCNTQVTQTVDFSLDPTFTIETTVPQDAVRSQRRTGRGRMGIYRFVTPG	480
Db	1446	GFDSVIDCNTQVTQTVDFSLDPTFTIETTLPQDAVRSQRRTGRGRMGIYRFVAPG	1505
QY	481	ERPSGMPDSSVLCEDYDAGCAWELTPAETSURLRAYLNTPLPVCOHLEFVESVETGL	540
Db	1506	ERPSGMPDSSVLCEDYDAGCAWELTPAETVURLRAYMNTPLPVCOHLEFVEGVEFTGL	1565
QY	541	THIDAFHLSQTKAGDNFPYLVAYQAVCARAQAPPSQWQMKLRLKPTLHGPTPLL	600
Db	1566	THIDAFHLSQTKSGENFPYLVAYQAVCARAQAPPSQWQMKLRLKPTLHGPTPLL	1625
QY	601	YRLGAVQNEVTTHTPIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTSQSVTVGR	660
Db	1626	YRLGAVQNEVTLTHPIKYIMTOMCSDLEVVSTWLVGGVLAALAAAYCLSTGCWIVGR	1685
QY	661	IILSGKPAIIPDRHVLRFEDMEECASHLPYIQOQGLAEQFKQKAIQLIQTATKQEA	720
Db	1686	IIVLSGKPAIIPDRHVLRFEDMEECASHLPYIQOQGLAEQFKQKAIQLIQTASQAEV	1745
QY	721	AAPVVEKSWTLEAFWAKHWNFTSGIYLAGLSTLPGNPAISLMAFTASITSLPTQH	780
Db	1746	ITPAVQTNQKLEVFNAKHMNFISGQYLAGLSTLPGNPAISLMAFTAAVTSPLTQ	1805
QY	781	TLFNILGGMVAAQAPPSAASAFVAGAGIAGAAVSGI GLGKLVLDIILAGYGAGVAGALVA	840
Db	1806	TLFNILGGMVAAQALAAPGAATAFVAGAGIAGAAVSGI GLGKLVLDIILAGYGAGVAGALVA	1865
QY	841	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR	900
Db	1866	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR	1925
QY	901	GNHVSPTHYPESDAAARVTOILSSITITOLLKELHOWINEDCTPCSGSLRDLWDWIC	960
Db	1926	GNHVSPTHYPESDAAARVTAISSLITVQLRLHOMISECTTPCSGSLRDLWDWIC	1985
QY	961	TVLTFDKTWLQSKLLPRLPGVPFSPQYKGVWRGDIMOTTCPGQAITGHVKNQSMR	1020
Db	1986	EVLSDFKTWLQSKLLPRLPGVPFSPQYKGVWRGDIMOTRCHCGAEITGHVKNQSMR	2045
QY	1021	IVGPRCTSNWHTGTFPINAYTTGCTPSPAPNYSRALWRVAEEVEVTRVGDFFHYVTGM	1080
Db	2046	IVGPRCTSNWHTGTFPINAYTTGCTPLPAPNYFALWRVAEEVEVTRVGDFFHYVTGM	2105
QY	1081	TTDNVPCPCQVPAPEFFTEVDGVRHLRYAPACKLLREEVTFVLVGLNQYLVGSQLPCEPE	1140
Db	2106	TTDNVPCPCQVPAPEFFTEVDGVRHLRYAPACKLLREEVTFVLVGLNQYLVGSQLPCEPE	2165
QY	1141	PDVAVLTSMLTDPSPHITATETAKRRLARGSPPSLASSASQLSAPSLKATCTTRHDSPPAD	1200
Db	2166	PDVAVLTSMLTDPSPHITATETAKRRLARGSPPSLASSASQLSAPSLKATCTANHDSPPAE	2225

Search completed: December 8, 2004, 12:36:07  
Job time : 138 secs

QY	1201	LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSFAEILRRSRKFFPRAM	1260
Db	2226	LIEANLLWQEMGNITRVESENKVVILDSFDPLVAEEDEREVSFAEILRRSRPARAL	2285
QY	1261	PIWARDYNPPLESKWDPDYVPPVVGCELPPAKAPPIPPRRKKTIVLSESTVSALA	1320
Db	2286	PWARDYNPPLEVETWKPDYVPPVVGCELPPPRPPPPPPRRKKTIVLTESTLSTALA	2345
QY	1321	ELATKTFGSESSEANDSGTATASPDQSDG-DAGSDVESYSMPLEGGPDPDLSDGS	1379
Db	2346	ELATKFGSSSTSGI-TGDNITTSBPAPSGCPSDSDVESYSMPLEGGPDPDLSDGS	2404
QY	1380	WSTVSSEA-SEDVVCCSMSTWTGALITPCAAEBETKLPIINALSNLRRHNLVYATTSRS	1438
Db	2405	WSTVSSGADTEDVCCSMSTWTGALVTPCAAEEQKLPIINALSNLRRHNLVYATTSRS	2464
QY	1439	ASLRQKKTVDRLQVLDHVRDVLKEMKAKASTVKAKLVSBEACKLTPPHSARSKFGYG	1498
Db	2465	ACQRQKKTVDRLQVLDHVRDVLKEMKAKASTVKAKLVSBEACKLTPPHSARSKFGYG	2524
QY	1499	AKDVRNLSKAVNHISVMKDLLEDTEPTDITIMAKNEVFCVQPEKGGKRPARLIVFPD	1558
Db	2525	AKDVRCHARKAVAHINSVMKDLLEDSTPTDITIMAKNEVFCVQPEKGGKRPARLIVFPD	2584
QY	1559	LGVVCEKALYDWTSLPOAVMGSSYGFQYSPQQRVEFLVNAWKAKKCPMGFAYDTRCF	1618
Db	2585	LGVVCEKALYDVVSKLPLAVMGSSYGFQYSPQQRVEFLVNAWKAKKCPMGFAYDTRCF	2644
QY	1619	DSTVTENDIRVESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKQNGCYRRCRASGV	1678
Db	2645	DSTVTESDIRTEALYQCCDLDPQARVAIKSLTERLYIGGPLTNSKQNGCYRRCRASGV	2704
QY	1679	LTTSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMT	1738
Db	2705	LTTSCGNTLTCLYKAAACRAAGLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMT	2764
QY	1739	RYSAAPPDGPPEYDLELITSCSSNVSVAHDAKGRVYVLTTRDPTTPLARAAWETARHTP	1798
Db	2765	RYSAAPPDGPPEYDLELITSCSSNVSVAHDAKGRVYVLTTRDPTTPLARAAWETARHTP	2824
QY	1799	VNSWLGNIIMYAPTLWARMILMTHPFSILLAQOLEKALDCQIYGACYSIEPLDLPPIQ	1858
Db	2825	VNSWLGNIIMFAPTLWARMILMTHPFSILLAQOLEKALDCQIYGACYSIEPLDLPPIQ	2884
QY	1859	RLHGLSAFSLHSPGSEINRVAASCLRKLGVPPLEVRHRRARSVRARLLSOGGRAATCGKY	1918
Db	2885	RLHGLSAFSLHSPGSEINRVAASCLRKLGVPPLEVRHRRARSVRARLLSOGGRAATCGKY	2944
QY	1919	LFNVAVRTKLKLTPIPAASQDLSSVFAVSGSDIYHLSRARPRWFMWCLLLSVGVG	1978
Db	2945	LFNVAVRTKLKLTPIPAAGRLDLSGWFTAGVSGSDIYHLSRARPRWFMWCLLLSVGVG	3004
QY	1979	IYLLPNR 1985	
Db	3005	IYLLPNR 3011	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:19:47 ; Search time 61 Seconds  
(without alignments)  
3130.988 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQQRGLLGCIIT.....FMWCLLLLSVGVIGIYLLPNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10170	97.2	3010	1 GNVVTC	genome polyprotein
2	10159	97.1	3010	1 A45573	genome polyprotein
3	10109	96.6	3010	1 GNVVJC	genome polyprotein
4	10069	96.2	3010	1 S18030	genome polyprotein
5	10066	96.2	3010	1 GNVVTV	genome polyprotein
6	9417.5	90.0	3011	1 GNVVVC	genome polyprotein
7	9398.5	89.8	3011	1 S40770	genome polyprotein
8	9266.5	88.5	3011	1 GNVVCH	genome polyprotein
9	8412.5	80.4	3014	1 JC5620	genome polyprotein
10	8037.5	76.8	3033	1 JQ1303	genome polyprotein
11	7959.5	76.1	3033	1 GNVVJ8	genome polyprotein
12	2383	22.8	492	2 PS0326	genome polyprotein - hepa
13	2098	20.0	876	2 PC2219	polyprotein - hepa
14	2078.5	19.9	3005	2 T08841	polyprotein - hepa
15	1980.5	18.9	2970	2 T08839	polyprotein - marm
16	1921.5	18.4	386	2 S68016	ATPase/RNA helicas
17	1855	17.7	365	2 J00879	NS5 protein - hepa
18	1755	16.8	365	2 J00880	NS5 protein - hepa
19	1555	14.9	716	2 JQ1366	polyprotein - hepa
20	1468	14.0	874	2 J00883	genome polyprotein
21	1457	13.9	874	2 J00881	genome polyprotein
22	1112	10.6	200	2 PQ0246	polyprotein (clone
23	1097	10.5	200	2 PQ0245	polyprotein (clone
24	1091	10.4	216	2 S21337	genome polyprotein
25	1048.5	10.0	1435	2 T01075	polyprotein - hepa
26	1036	9.9	194	2 A54317	probable nonstruct
27	1034	9.9	194	2 S06067	nonstructural prot
28	986	9.4	200	2 PQ0248	polyprotein (clone
29	983	9.4	184	2 A61196	genome polyprotein

#### ALIGNMENTS

##### RESULT 1

GNWVTC

genome polyprotein - hepatitis C virus

N:Contains: capsid protein C; envelope protein M; hepatitis M; hepatitis (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C:Accession: A38465

R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J. Virol. 65, 1105-1113, 1991

A:Title: Structure and organization of the hepatitis C virus genome isolated from human c

A:Reference number: A38465; MUID:91140698; PMID:1847440

A:Accession: A38465

A:Molecule type: Genomic RNA

A:Residues: 1-3010 <TAK>

A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PID:AAA72945.1; PID:g32977

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <BPM>

F:192-389/Product: major envelope protein E #status predicted <MEE>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitis C virus genome polyprotein

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: DEHX motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,224

##### Query Match

Best Local Similarity 97.2%; Score 10170; DB 1; Length 3010;

Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MAPITAYSQQRGLLGCIITSLTGRDNQVEGVVSTATQSFATCVCNGVCTVYHGA 60

Db 1026 LAPITAYSQQRGLLGCIITSLTGRDNQVEGVVSTATQSFATCVCNGVCTVYHGA 1085

Qy 61 GSKTLAKPGKITOMYTNVDQDLVGVQAPPGARSLTPTCTGSSDLVLTTRHADVPIVRRR 120

Db 1086 GSKTLAKPGKITOMYTNVDQDLVGVQAPPGARSLTPTCTGSSDLVLTTRHADVPIVRRR 1145

Qy 121 GDSRGSLLSRPVSYLKSGSGPCLLCPSGHAVGIFRAVCTRGVAKAVDFVPVESMETTM 180

Db 1146 GDSRGSLLSRPVSYLKSGSGPCLLCPSGHAVGIFRAVCTRGVAKAVDFVPVESMETTM 1205

Qy 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKGSKTKVPAAYAAQGYKVLVLPNSVAATLGF 240

Db 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKGSKTKVPAAYAAQGYKVLVLPNSVAATLGF 1265

QY 241 AYMSKAGIDPNIRTCVRIITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
Db 1266 AYMSKAGIDPNIRTCVRIITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 1325  
QY 301 ILGIGTVLDOAETAGARLVVLAATATPPGTVVPHPNIEEVALSGTGEIPYKAIPIETI 360  
Db 1326 ILGIGTVLDOAETAGARLVVLAATATPPGTVVPHPNIEEVALSGTGEIPYKAIPIEA 1385  
QY 361 KGRHILIFCHSKKKDELAAKLSGLGNVAVYVYRGLDVSVIPTSGDVVIVATDALTMTGFT 420  
Db 1386 KGRHILIFCHSKKKDELAAKLSGLGNVAVYVYRGLDVSVIPTSGDVVIVATDALTMTGFT 1445  
QY 421 GPDFSVLDNCTVQTVDFSLDFTETITTTVPQDAVRSORRGTRGRGMGIYRFVTPG 480  
Db 1446 GPDFSVLDNCTVQTVDFSLDFTETITTTVPQDAVRSORRGTRGRGMGIYRFVTPG 1505  
QY 481 ERPSGMFDSVLCBCEYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVFTGL 540  
Db 1506 ERPSGMFDSVLCBCEYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVFTGL 1565  
QY 541 THIDAHFLSOTKAGDNFPVLYAYQATVCAQAAPPSPWDQMKCLIRLKPRTLHGPTPLL 600  
Db 1566 THIDAHFLSOTKAGDNFPVLYAYQATVCAQAAPPSPWDQMKCLIRLKPRTLHGPTPLL 1625  
QY 601 YRLGAVQNEVTTTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVIVGR 660  
Db 1626 YRLGAVQNEVTTTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVIVGR 1685  
QY 661 IILSGPAITPDREVLRYEDEMEECASHLPYIEQGMQLAEQFKQKAGILLOTATKQAEA 720  
Db 1686 IILSGPAITPDREVLRYEDEMEECASHLPYIEQGMQLAEQFKQKAGILLOTATKQAEA 1745  
QY 721 AAPVWESKRTLEAFNAKHWNNFISGIQYLAGUSTLPGNPAIASLMAFTASITSPLTQH 780  
Db 1746 AAPVWESKRTLEAFNAKHWNNFISGIQYLAGUSTLPGNPAIASLMAFTASITSPLTQH 1805  
QY 781 TLLFNILGGWVAQALAPSAASAVGAGIAGAAVSGTGLGKVLVDIILAGYAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQALAPSAASAVGAGIAGAAVSGTGLGKVLVDIILAGYAGVAGALVA 1865  
QY 841 FKVMSGEMSPEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNNRLIAPASR 900  
Db 1866 FKVMSGEMSPEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNNRLIAPASR 1925  
QY 901 GNHVSPTHYVPESDAARVTOILSSLIITQLKRLHOWINEDCSTPCSGSLRVDWDMIC 960  
Db 1926 GNHVSPTHYVPESDAARVTOILSSLIITQLKRLHOWINEDCSTPCSGSLRVDWDMIC 1985  
QY 961 TVLTDFKTLQSKLLPRLPGVPFPCORGYKGVWRGDGIMOTTCPCGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFKTLQSKLLPRLPGVPFPCORGYKGVWRGDGIMOTTCPCGAQITGHVKNQSMR 2045  
QY 1021 IGVPRTCSTNWHGTFPINAYTTGCTPSPAPNYSRALWRVAEEYEVTRVGDPHYVTGM 1080  
Db 2046 IGVPRTCSTNWHGTFPINAYTTGCTPSPAPNYSRALWRVAEEYEVTRVGDPHYVTGM 2105  
QY 1081 TTDNVKPCQVPAPEFTEVDGVLRLHYAPACKPDLLEBEVTFVLGILNQYLVSQIPEPE 1140  
Db 2106 TTDNVKPCQVPAPEFTEVDGVLRLHYAPACKPDLLEBEVTFVLGILNQYLVSQIPEPE 2165  
QY 1141 PDVAVLTSMLTDPDHSHTAETAKRLRARGSPPSLASSASQSLASPLKATCTTRHDSPPAD 1200  
Db 2166 PDVAVLTSMLTDPDHSHTAETAKRLRARGSPPSLASSASQSLASPLKATCTTRHDSPPAD 2225  
QY 1201 LIEANLLWRQBMGNNITRVSENNKVILDSFEPIQAEDEDEREVSVPAELRRSRKFPAM 1260  
Db 2226 LIEANLLWRQBMGNNITRVSENNKVILDSFEPIQAEDEDEREVSVPAELRRSRKFPAM 2285  
QY 1261 PIWARPDPNPLLESWKDPDVPVPPVHVCPLPPAKAPPIPPRRKRTVVLSESVSSALA 1320  
Db 2286 PIWARPDPNPLLESWKDPDVPVPPVHVCPLPPAKAPPIPPRRKRTVVLSESVSSALA 2345  
QY 1321 ELATKTFGSSSAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLGEPEGDPDLSGWS 1380

Db 2346 ELATKTFGSSSAVDSGTATALPDQASDDGDKGSDVESYSSMPPLGEPEGDPDLSGWS 2405  
QY 1381 STVSEASBDDVCCSMYSYWTGALITPCAAEETKLPINALNSLLRHNNLVYATTSBAS 1440  
Db 2406 STVSEASBDDVCCSMYSYWTGALITPCAAEESKLPINALNSLLRHNNLVYATTSBAS 2465  
QY 1441 LRQKKYTFORLOVLDHHDYDLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
Db 2466 LRQKKYTFORLOVLDHHDYDLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
QY 1501 DVNLSKSAVNHRISVWVKDILETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPPDLG 1560  
Db 2526 DVNLSKSAVNHRISVWVKDILETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPPDLG 2585  
QY 1561 VRVCEKMALYDVVSTLTPQAVMGSSYGFQYSPGQRFVFLYNAMKAKKCPMGAFVDTTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLTPQAVMGSSYGFQYSPGQRFVFLYNAMKAKKCPMGAFVDTTRCFDS 2645  
QY 1621 TVTENDIRVEESYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNCGYRRCRASGVLT 1680  
Db 2646 TVTENDIRVEESYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNCGYRRCRASGVLT 2705  
QY 1681 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEDEASLRAPTEAMTRY 1740  
Db 2706 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEDEASLRAPTEAMTRY 2765  
QY 1741 SAPPGPPPKPEYDLELITSCSSNVSVNAHDASGRVYVYLTDPDPTPLARAWETARHTPVN 1800  
Db 2766 SAPPGPPPKPEYDLELITSCSSNVSVNAHDASGRVYVYLTDPDPTPLARAWETARHTPVN 2825  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAOEKLKALDCQIYACYSIEPLDLPQIITQRL 1860  
Db 2826 SWLGNIIIMYAPTILWARMILMTHFFSILLAOEKLKALDCQIYACYSIEPLDLPQIITQRL 2885  
QY 1861 HGLSARFSLHSYSGEINRVASCLRLKGLVPPPLVRHRRARSVRARLLSQGGRAATCGKYLE 1920  
Db 2886 HGLSARFSLHSYSGEINRVASCLRLKGLVPPPLVRHRRARSVRARLLSQGGRAATCGKYLE 2945  
QY 1921 NNAVRYTKLUTPAPASQDLSSWFVAGYSGGDIYHLSLRARPRFWFNCILLLSVGVGIY 1980  
Db 2946 NNAVRYTKLUTPAPASQDLSSWFVAGYSGGDIYHLSLRARPRFWFNCILLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

## RESULT 2

A45573

genome polyprotein - hepatitis C virus (strain JT)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a); nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C:Accession: A45573

P:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,

Virus Res. 23, 39-53, 1992

A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier:

A:Reference number: A45573; MUID:92295714; PMID:1318627

A:Accession: A45573

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-3010 &lt;TAN&gt;

A:Cross-references: UNIPROT\_Q00269; GB:D11168; GB:D01171; NID:G221612; PIDN:BAA01943.1.7

A:Experimental source: HCV-JT

C:Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; seri

F:2-115/Product: capsid protein C #status predicted &lt;CPC&gt;

F:116-191/Product: envelope protein M #status predicted &lt;EPM&gt;

F:192-389/Product: major envelope protein E #status predicted &lt;MEE&gt;

F:390-729/Product: nonstructural protein NS1 #status predicted &lt;NS1&gt;



Db 3006 LLENR 3010

RESULT 3  
GNWVCJ  
genome polyprotein - hepatitis C virus (strain J)  
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: A39253; PMID:91088550; PMID:2175903  
R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimotoh  
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990  
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients  
A:Reference number: A39253; MUID:91088550; PMID:2175903  
A:Accession: A39253  
A:Molecule type: genomic RNA  
A:Residues: 1-3010 <XA2>  
A:Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611  
R:Kato, N.; Ohkoshi, S.; Shimotohno, K.  
Proc. Jpn. Acad. 65B, 219-223, 1989  
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari  
A:Reference number: PS0085  
A:Accession: PS0086  
A:Molecule type: genomic RNA  
A:Residues: 2850-2707 <XA2>  
A:Experimental source: Japanese isolate  
C:Comment: The cleavage sites of this polyprotein have not been determined.  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: Atp; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin  
F:2-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEE>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis C virus #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Query Match 96.6%; Score 10109; DB 1; Length 3010;  
Best Local Similarity 95.6%; Pred. No. 0;  
Matches 1896; Conservative 50; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MAPITAYSOQTRGLGCIITSLTGRDNOVEGVQVYSTATQSFATCVNGVCMVTVHGA 60  
Db 1026 LAPITAYSOQTRGLGCIITSLTGRDKNQVDGQVLSLATQSFATCVNGVCMVTVHGA 1085

Qy 61 GSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTCTCGSSDLVLTTRHADYIPVRRR 120  
Db 1086 GSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSMTCTCGSSDLVLTTRHADYIPVRRR 1145

Qy 121 GDSRGSLLSPRPVSYLKGSSGGLPLCPSSHAVGIFRAVCTRGVAKAVDFVPVSMETTM 180  
Db 1146 GDSRGSLLSPRPVSYLKGSSGGLPLCPSSHVGVIFRAVCTRGVAKAVDFIPVSMETTM 1205

Qy 181 RSPVFTDSSPPVAPQFQVLAHLHAPTGSKGSKTKVPAAYAAQYKVLVLPNSVAATLGF 240  
Db 1206 RSPVFTDSSPPVAPQFQVLAHLHAPTGSKGSKTKVPAAYAAQYKVLVLPNSVAATLGF 1255

Qy 241 AYMSKAHGIDPNINTGVRTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 1266 AYMSKAHGIBENIRTVRTITGPIITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 1325

Qy 301 ILGIGTVLDAQETAGARLVLATPPGTVTPHPNIEVALSTGRIPIFYGKAIPETI 360  
Db 1326 ILGIGTVLDAQETAGARLVLATPPGTVTPHPNIEVALSTGRIPIFYGKAIPETI 1385

Qy 361 KGRHILIFCHSKKCDLAAKLSGLNVAAYVRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Db 1386 KGRHILIFCHSKKCDLAAKLSGLNVAAYVRGLDVSIVPTSGDVIIVATDALMTGFT 1445

Qy 421 GDFDSVIDCNTCTVTQVDFSLDPTFTTETTPQDAVSRRGRGRTGRGMGYRFTVTPG 480  
Db 1446 GDFDSVIDCNTCTVTQVDFSLDPTFTTETTPQDAVSRAQRGRGRTGRGMGYRFTVTPG 1505

Qy 481 ERSGMFDSVLCVCEYDAGCAWELTAPETSRLRAYLNTPLGVPVCDHLEFESVETGL 540  
Db 1506 ERSGMFDSVLCVCEYDAGCAWELTAPETSRLRAYLNTPLGVPVCDHLEFESVETGL 1565

Qy 541 THIDAFSLQTKAGDNFYLVAQATVCARAQAPPSWDMQMKCLIRLKLPTLHGPTPL 600  
Db 1566 THIDAFSLQTKAGDNFYLVAQATVCARAQAPPSWDMQMKCLIRLKLPTLHGPTPL 1625

Qy 601 YRLGAVQNEVTTHTPTIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTGSVTVGR 660  
Db 1626 YRLGAVQNEVTTHTPTIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTGTGSVTVGR 1685

Qy 661 IILSGKPAIIPDREVLYRBEDEECASHLPYIEQGMQLAEQFKQKALGLQATKQAEA 720  
Db 1686 IILSGKPAIIPDREVLYRBEDEECASHLPYIEQGMQLAEQFKQKALGLQATKQAEA 1745

Qy 721 AAPVBSKWRTEAFWAKHWNFIISGIQYLAGLSTLPNPAIASLMAFTASITSLTQ 780  
Db 1746 AAPVBSKWRTEAFWAKHWNFIISGIQYLAGLSTLPNPAIASLMAFTASITSLTQ 1805

Qy 781 TLLFNILGGWVAQAAPPASAASAVGAGIAGAAGVSGIGLKVLIIDILAGYAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQAAPPASAASAVGAGIAGAAGVSGIGLKVLIIDILAGYAGVAGALVA 1865

Qy 841 FKVMGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAVQWNNRLIAFASR 900  
Db 1866 FKVMGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEAVQWNNRLIAFASR 1925

Qy 901 GNVSPHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCSTPCSGSLRDVMDWIC 960  
Db 1926 GNVSPHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCSTPCSGSLRDVMDWIC 1985

Qy 961 TVLTDFKWLQSKLLPRLPGVPFFSCQYKGVWVGDIQMTTCPCGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFKWLQSKLLPRLPGVPFFSCQYKGVWVGDIQMTTCPCGAQITGHVKNQSMR 2045

Qy 1021 IVGPRTCSNTHGTFEPINATYTCPTSPAPNYSRALRWAAEYEVTVGDHYVTGM 1080  
Db 2046 IVGPRTCSNTHGTFEPINATYTCPTSPAPNYSRALRWAAEYEVTVGDHYVTGM 2105

Qy 1081 TTDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE 1140  
Db 2106 TTDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE 2165

Qy 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSPASSASQISAPSLKATCTTRHDSPPAD 1200  
Db 2166 PDVAVLTSMLTDSHITAEAKRLARGSPSPASSASQISAPSLKATCTTRHDSPPAD 2225

Qy 1201 LIENALLWRQEMGNITRVSEKNVILDSFELQAEDEEREVSVPAAEILRRSRKFFPRAM 1260  
Db 2226 LIENALLWRQEMGNITRVSEKNVILDSFELQAEDEEREVSVPAAEILRRSRKFFPAL 2285

Qy 1261 PIWARPYNPPLLESWKDDPDYFPVHVHGCPLPAKAPPIPPRRKRVTVLSSESTVSALA 1320  
Db 2286 PIWARPYNPPLLESWKDDPDYFPVHVHGCPLPAKAPPIPPRRKRVTVLSSESTVSALA 2345

Qy 1321 ELATKTFGSSSNAVDSGTATASDPQSDGSDGSDVESYSSMPPLGEPCGDPDLSGWSW 1380  
Db 2346 ELATKTFGSSSNAVDSGTATASDPQSDGSDGSDVESYSSMPPLGEPCGDPDLSGWSW 2405

Qy 1381 STVSEBASEDVCCMSYTWGALITPCAABETKLPINALSNSLLRHNLVYATTSRSAS 1440  
Db 2406 STVSEBASEDVCCMSYTWGALITPCAABETKLPINALSNSLLRHNLVYATTSRSAS 2465

Qy 1441 LRQKVTDFRLQVLDDHYRDVLKEMKAKASTVKAKLLSVBEACKLTTPHSARSKEGYAK 1500



```
Db 2466 LQKKVTEDRLQVLDHRYDLVKEMKAKASTVKARLLSIEACKLTPPHSAKSKFEGYAK 2525
Qy 1501 DVRLSSKAVNHRSVWMDLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560
Db 2526 DVRLSSKAVNHRSVWMDLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 2585
Qy 1561 VRVCEKMALYDVVSTLPOAVNGSSYGFQYSPQQRVEFLVNAWAKKCPMGFAYDTRCFDS 1620
Db 2586 VRVCEKMALYDVVSTLPOAVNGSPSYGFQYSPQQRVEFLVNTWKSCKPMGFSDYDTRCFDS 2645
Qy 1621 TVTENDIRVESIYOCCDLAPEARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLT 1680
Db 2646 TVTENDIRTESIYOCCDLAPEARQAIRSLTERLYVGGPLTNSKQNGCYRRCRASGVLT 2705
Qy 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDGLVVICESAGTQDEASLRAFTAMTRY 1740
Db 2706 TSCGNLTTCYLKATAACRAAKLQDCTMLVNGDGLVVICESAGTQDEAALRAFTAMTRY 2765
Qy 1741 SAPGDPKPEYDUELITSCSSNVSAHDAGSKRVVYLTRDPTTPPARAAWETARHTPVN 1800
Db 2766 SAPGDPKPEYDUELITSCSSNVSAHDAGSKRVVYLTRDPTTPPARAAWETARHTPVN 2825
Qy 1801 SWLGNIMYAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPLDLPQIOLRL 1860
Db 2826 SWLGNIMYAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPLDLPQIOLRL 2885
Qy 1861 HGLSAFSLHSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPG 1920
Db 2886 HGLSAFSLHSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPG 2945
Qy 1921 NWAHTKILKLTPIPAASOLDIJSWFWAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY 1980
Db 2946 NWAHTKILKLTPIPAASOLDIJSWFWAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY 3005
Qy 1981 LLEPNR 1985
Db 3006 LLEPNR 3010

RESULT 4
S18030
Genome polyprotein - hepatitis C virus (isolate JK1)
N:contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a); nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Variety: isolate JK1
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:G59478; PIDN:CAAA3793.1; PID:G59479
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA
A:Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>
A:Cross-references: EMBL:X61591
A:Note: this sequence is inconsistent with the nucleotide translation
A:Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
A:Note: sequence extracted from NCBI backbone (NCBI:121747, NCBI:121748)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
P:2-115/Product: capsid protein C #status predicted <CPC>
```

F:116-131/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEE>  
F:730-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1007-1615/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1230-1237/Product: hepatitis C virus #status predicted <NS3>  
F:1312-1317/Region: nucleotide-binding motif A (P-loop)  
F:1316-1319/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1863-2010/Product: nonstructural protein NS4a #status predicted <N4A>  
F:1863-2010/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (Asn)

Query Match 96.2%; Score 10069; DB 1; Length 3010;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 1901; Conservative 39; Mismatches 45; Indels 0; Gaps 0;

```
Qy 1 MAPITAYSOOTRGLGGLCIIITSLTGRDRNQVGEVQVSTATQSEFLATCVNGVCTVYHGA 60
Db 1026 LAPITAYSQQRGLFGCIIVTSLTGRDRNQVGEVQVSTATQSEFLATCVNGVCTVYHGA 1085
Qy 61 GSKTLAGEPKGITQMYTNVDQDLVGMQAPPGARSILPTCTCGSSDLYLVTRHADVIPVRRR 120
Db 1086 GSKTLAGEPKGIPINQMYTNVDQDLVGMQAPSGAASILPTCTYSSDLYLVTRHADVIPVRRR 1145
Qy 121 GDSRGSLLSPRPVSYLVKSSGGPPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVSEMETTM 180
Db 1146 GDSRGSLLSPRPVSYLVKSSGGPPLLCPGSHAVGIFPRAAVCTRGVAKAVDFIPVSEMETTM 1205
Qy 181 RSPVFTDNSSPPAVPQTEFQVAHLHAPTGSKSTKVPAAAYAGYKVLVNPVAAATLFG 240
Db 1206 RSPVFTDNSSPPAVPQTEFQVAHLHAPTGSKSTKVPAAAYAGYKVLVNPVAAATLFG 1265
Qy 241 AYMSKAHGDINIRGTITGAPITVYSTYKFLADGCGSGGAYDIIICDECHSDSTT 300
Db 1266 AYMSKAHGDINIRGTITGAPITVYSTYKFLADGCGSGGAYDIIICDECHSDSTT 1325
Qy 301 ILGIGTVLDQAEAGARLVVLAATPQSGVTVPHNIEVALSSTGEIPFGYKALPIETI 360
Db 1326 ILGIGTVLDQAEAGARLVVLAATPQSGVTVPHNIEVALSSTGEIPFGYKALPIETI 1385
Qy 361 KGRHLIFCHSKKKDELAALKSLGLNNAVAYVGLDVSVIPTSGDVIIVVATDALMTGFT 420
Db 1386 KGRHLIFCHSKKKDELAALKSLGLNNAVAYVGLDVSVIPTSGDVIIVVATDALMTGFT 1445
Qy 421 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSORRGRTGRGRGIYFVTPG 480
Db 1446 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSORRGRTGRGRGIYFVTPG 1505
Qy 481 ERPSGMFSSVLCBCEYDAGCAWYELTPAETSVRRLAYLNTFGLPVCQDHLFWESVFTGL 540
Db 1506 ERPSGMFSSVLCBCEYDAGCAWYELTPAETSVRRLAYLNTFGLPVCQDHLFWESVFTGL 1565
Qy 541 THIDAHFLSQTQKQADNPFLVAVQATVCARAQAPPPSDQMWKCLIRLKLTHGHTPLL 600
Db 1566 THIDAHFLSQTQKQADNPFLVAVQATVCARAQAPPPSDQMWKCLIRLKLTHGHTPLL 1625
Qy 601 YRLGAVQNEVTTTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR 660
Db 1626 YRLGAVQNEVTLTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR 1685
Qy 661 IILSGKPAIIPDREVLVREPFDEMEECASHLPIYISGMOLARQFKQKALGLLTATKQAEA 720
Db 1686 IILSGKPAIIPDREVLVREPFDEMEECASHLPIYISGMOLARQFKQKALGLLTATKQAEA 1745
Qy 721 AAPVVSQWRTLEAFWAKHWNFTSGIYLAGLSTLPCNPALASIMAFSTSTPLTQCH 780
Db 1746 AAPVVSQWRTLEAFWAKHWNFTSGIYLAGLSTLPCNPALASIMAFSTSTPLTQCH 1805
Qy 781 TLLFNILGSGWAAQLAPPASAFYAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 840
Db 1806 TLLFNILGSGWAAQLAPPASAFYAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 1865
```

QY 841 FKVMGEMPTEDLVNLLPALLSPGALVGVVCAAILRRHVGPCEGAVQWNRLLIAPSR 900  
Db 1866 FKMGSEMPSTEDLVNLLPALLSPGALVGVVCAAILRRHVGPCEGAVQWNRLLIAPSR 1925  
QY 901 GNVSPHYVPESDAARVTKILSSLLITQLKRLHWINEDCSTPCSGMLRDVMDWIC 960  
Db 1926 GNVSPHYVPESDAARVTKILSSLLITQLKRLHWINEDCSTPCSGMLRDVMDWIC 1985  
QY 961 TVLTDFTXWLSKLLPLPGLPVPFPSCQRYGKGVWRGDMGIMQTTCCPCGAQITGHVKGSMR 1020  
Db 1986 TVLTDFTXWLSKLLPLPGLPVPFPSCQRYGKGVWRGDMGIMQTTCCPCGAQITGHVKGSMR 2045  
QY 1021 IVPBRTCSNTHGTFPFINAVTTGCTSPAPNYSRALWRVAAEYVSVTRVGDHYVTGM 1080  
Db 2046 IVPBRTCSNTHGTFPFINAVTTGCTSPAPNYSRALWRVAAEYVSVTRVGDHYVTGM 2105  
QY 1081 TTDNVKPCQVPAPEFFTEVDGVLHRYAPACKPLLRDEEVTFVLGNQLYVLSGSLPCEPE 1140  
Db 2106 TTDNVKPCQVPAPEFFTEVDGVLHRYAPACKPLLRDEEVTFVLGNQLYVLSGSLPCEPE 2165  
QY 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Db 2166 PDVTVLTSMLTDSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWROEMGMGNITRVESENKVVILDSFEPLOAEDEDEREVGPAAEILRRSRKFFPRAM 1260  
Db 2226 LIEANLLWROEMGMGNITRVESENKVVILDSFEPLOAEDEDEREVGPAAEILRRSRKFFPAL 2285  
QY 1261 PIWARPYNPLLESWKDDPYVPVPHVHCPLPPAKAPPPIPPRRKRTVWLSSESTVSALA 1320  
Db 2286 PIWARPYNPLLESWKDDPYVPVPHVHCPLPPAKAPPPIPPRRKRTVWLSSESTVSALA 2345  
QY 1321 ELATKTPGSESSAVDGTATSDQSDGDDAGSDVESYSSMPLPGEFGDPLDGSW 1380  
Db 2346 ELATKTPGSESSAVDGTATSDQSDGDDAGSDVESYSSMPLPGEFGDPLDGSW 2405  
QY 1381 STVSEERASEDVCCSMSTYTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSAS 1440  
Db 2406 STVSEERASEDVCCSMSTYTGALITPCAAEESKLPINPLSNSLLRHHNLVYATTSAS 2465  
QY 1441 LRQKVTFRDRLQVLDHYRDVLEKMKAKASTVAKALLSVEACKLTPPHSARSFGYGA 1500  
Db 2466 LRQKVTFRDRLQVLDHYRDVLEKMKAKASTVAKALLSVEACKLTPPHSARSFGYGA 2525  
QY 1501 DVRLSSKANVHRSVMKLLDTEETDITDITMAKNEVFCVQPEKGRKPARLTVFPDLG 1560  
Db 2526 DVRLSSKANVHRSVMKLLDTEETDITDITMAKNEVFCVQPEKGRKPARLTVFPDLG 2585  
QY 1561 VRCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
Db 2586 VRCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 2645  
QY 1621 TVTENDIRVEESYQCCDLAPARQARSUTERLYIGGPLTNSKGQCYRRCRASGVL 1680  
Db 2646 TVTENDIRVEESYQCCDLAPARQARSUTERLYIGGPLTNSKGQCYRRCRASGVL 2705  
QY 1681 TSCGNTLTCYLAACAAKALQDCTMLVCGDDLAVLCEAGTQDEASRAFTTEATRY 1740  
Db 2706 TSCGNTLTCYLAACAAKALQDCTMLVCGDDLAVLCEAGTQDEASRAFTTEATRY 2765  
QY 1741 SAPPGDPKPEYDLELITSCSNVSVVAHDASGRVYVYLTDRPTPLPAAWETARHTFPV 1800  
Db 2766 SAPPGDPKPEYDLELITSCSNVSVVAHDASGRVYVYLTDRPTPLPAAWETARHTFPV 2825  
QY 1801 SWLGNITMYAPTLLWARMILTHFSSILLAEQLEKALDCCQYGCYSIEBPLDLPQIITQRL 1860  
Db 2826 SWLGNITMYAPTLLWARMILTHFSSILLAEQLEKALDCCQYGCYSIEBPLDLPQIITQRL 2885  
QY 1861 HGLSAPSLHSYSGEINRVASCLRLKGVPPPLRVWRHARSVRARLLSOGGRAATCGKYL 1920  
Db 2886 HGLSAPSLHSYSGEINRVASCLRLKGVPPPLRVWRHARSVRARLLSOGGRAATCGKYL 2945  
QY 1921 NNAVRTKLTPTPAASQLDLSWFWAGYSGGDIYHLSLRARPRFWFWMCLLLSSVGVGIY 1980

Db 2946 NNAVRTKLTPTPAASQLDLSWFWAGYSGGDIYHLSLRARPRFWFWMCLLLSSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010  
RESULT 5  
GNVWTW  
genome polyprotein - hepatitis C virus (strain Taiwan)  
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain Taiwan)  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: A40244  
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.  
Virology 188, 102-113, 1992  
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the  
A:Reference number: A40244; MUID:92230206; PMID:1314449  
A:Molecule type: genomic RNA  
A:Residues: 1-3010 <CHE>  
A:Cross-references: UNIPROT:P29846; GB:M84754  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: AIP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
F:1-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEE>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,233,234,250,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2042,2043,2044,2045,2046,2047,2048,2049,2050,2051,2052,2053,2054,2055,2056,2057,2058,2059,2060,2061,2062,2063,2064,2065,2066,2067,2068,2069,2070,2071,2072,2073,2074,2075,2076,2077,2078,2079,2080,2081,2082,2083,2084,2085,2086,2087,2088,2089,2090,2091,2092,2093,2094,2095,2096,2097,2098,2099,2100,2101,2102,2103,2104,2105,2106,2107,2108,2109,2110,2111,2112,2113,2114,2115,2116,2117,2118,2119,2120,2121,2122,2123,2124,2125,2126,2127,2128,2129,2130,2131,2132,2133,2134,2135,2136,2137,2138,2139,2140,2141,2142,2143,2144,2145,2146,2147,2148,2149,2150,2151,2152,2153,2154,2155,2156,2157,2158,2159,2160,2161,2162,2163,2164,2165,2166,2167,2168,2169,2170,2171,2172,2173,2174,2175,2176,2177,2178,2179,2180,2181,2182,2183,2184,2185,2186,2187,2188,2189,2190,2191,2192,2193,2194,2195,2196,2197,2198,2199,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209,2210,2211,2212,2213,2214,2215,2216,2217,2218,2219,2220,2221,2222,2223,2224,2225,2226,2227,2228,2229,2230,2231,2232,2233,2234,2235,2236,2237,2238,2239,2240,2241,2242,2243,2244,2245,2246,2247,2248,2249,2250,2251,2252,2253,2254,2255,2256,2257,2258,2259,2260,2261,2262,2263,2264,2265,2266,2267,2268,2269,2270,2271,2272,2273,2274,2275,2276,2277,2278,2279,2280,2281,2282,2283,2284,2285,2286,2287,2288,2289,2290,2291,2292,2293,2294,2295,2296,2297,2298,2299,2300,2301,2302,2303,2304,2305,2306,2307,2308,2309,2310,2311,2312,2313,2314,2315,2316,2317,2318,2319,2320,2321,2322,2323,2324,2325,2326,2327,2328,2329,2330,2331,2332,2333,2334,2335,2336,2337,2338,2339,2340,2341,2342,2343,2344,2345,2346,2347,2348,2349,2350,2351,2352,2353,2354,2355,2356,2357,2358,2359,2360,2361,2362,2363,2364,2365,2366,2367,2368,2369,2370,2371,2372,2373,2374,2375,2376,2377,2378,2379,2380,2381,2382,2383,2384,2385,2386,2387,2388,2389,2390,2391,2392,2393,2394,2395,2396,2397,2398,2399,2400,2401,2402,2403,2404,2405,2406,2407,2408,2409,2410,2411,2412,2413,2414,2415,2416,2417,2418,2419,2420,2421,2422,2423,2424,2425,2426,2427,2428,2429,2430,2431,2432,2433,2434,2435,2436,2437,2438,2439,2440,2441,2442,2443,2444,2445,2446,2447,2448,2449,2450,2451,2452,2453,2454,2455,2456,2457,2458,2459,2460,2461,2462,2463,2464,2465,2466,2467,2468,2469,2470,2471,2472,2473,2474,2475,2476,2477,2478,2479,2480,2481,2482,2483,2484,2485,2486,2487,2488,2489,2490,2491,2492,2493,2494,2495,2496,2497,2498,2499,2500,2501,2502,2503,2504,2505,2506,2507,2508,2509,2510,2511,2512,2513,2514,2515,2516,2517,2518,2519,2520,2521,2522,2523,2524,2525,2526,2527,2528,2529,2530,2531,2532,2533,2534,2535,2536,2537,2538,2539,2540,2541,2542,2543,2544,2545,2546,2547,2548,2549,2550,2551,2552,2553,2554,2555,2556,2557,2558,2559,2560,2561,2562,2563,2564,2565,2566,2567,2568,2569,2570,2571,2572,2573,2574,2575,2576,2577,2578,2579,2580,2581,2582,2583,2584,2585,2586,2587,2588,2589,2590,2591,2592,2593,2594,2595,2596,2597,2598,2599,2600,2601,2602,2603,2604,2605,2606,2607,2608,2609,2610,2611,2612,2613,2614,2615,2616,2617,2618,2619,2620,2621,2622,2623,2624,2625,2626,2627,2628,2629,2630,2631,2632,2633,2634,2635,2636,2637,2638,2639,2640,2641,2642,2643,2644,2645,2646,2647,2648,2649,2650,2651,2652,2653,2654,2655,2656,2657,2658,2659,2660,2661,2662,2663,2664,2665,2666,2667,2668,2669,2670,2671,2672,2673,2674,2675,2676,2677,2678,2679,2680,2681,2682,2683,2684,2685,2686,2687,2688,2689,2690,2691,2692,2693,2694,2695,2696,2697,2698,2699,2700,2701,2702,2703,2704,2705,2706,2707,2708,2709,2710,2711,2712,2713,2714,2715,2716,2717,2718,2719,2720,2721,2722,2723,2724,2725,2726,2727,2728,2729,2730,2731,2732,2733,2734,2735,2736,2737,2738,2739,2740,2741,2742,2743,2744,2745,2746,2747,2748,2749,2750,2751,2752,2753,2754,2755,2756,2757,2758,2759,2760,2761,2762,2763,2764,2765,2766,2767,2768,2769,2770,2771,2772,2773,2774,2775,2776,2777,2778,2779,2780,2781,2782,2783,2784,2785,2786,2787,2788,2789,2790,2791,2792,2793,2794,2795,2796,2797,2798,2799,2800,2801,2802,2803,2804,2805,2806,2807,2808,2809,2810,2811,2812,2813,2814,2815,2816,2817,2818,2819,2820,2821,2822,2823,2824,2825,2826,2827,2828,2829,2830,2831,2832,2833,2834,2835,2836,2837,2838,2839,2840,2841,2842,2843,2844,2845,2846,2847,2848,2849,2850,2851,2852,2853,2854,2855,2856,2857,2858,2859,2860,2861,2862,2863,2864,2865,2866,2867,2868,2869,2870,2871,2872,2873,2874,2875,2876,2877,2878,2879,2880,2881,2882,2883,2884,2885,2886,2887,2888,2889,2890,2891,2892,2893,2894,2895,2896,2897,2898,2899,2900,2901,2902,2903,2904,2905,2906,2907,2908,2909,2910,2911,2912,2913,2914,2915,2916,2917,2918,2919,2920,2921,2922,2923,2924,2925,2926,2927,2928,2929,2930,2931,2932,2933,2934,2935,2936,2937,2938,2939,2940,2941,2942,2943,2944,2945,2946,2947,2948,2949,2950,2951,2952,2953,2954,2955,2956,2957,2958,2959,2960,2961,2962,2963,2964,2965,2966,2967,2968,2969,2970,2971,2972,2973,2974,2975,2976,2977,2978,2979,2980,2981,2982,2983,2984,2985,2986,2987,2988,2989,2990,2991,2992,2993,2994,2995,2996,2997,2998,2999,3000,3001,3002,3003,3004,3005,3006,3007,3008,3009,3010,3011,3012,3013,3014,3015,3016,3017,3018,3019,3020,3021,3022,3023,3024,3025,3026,3027,3028,3029,3030,3031,3032,3033,3034,3035,3036,3037,3038,3039,3040,3041,3042,3043,3044,3045,3046,3047,3048,3049,3050,3051,3052,3053,3054,3055,3056,3057,3058,3059,3060,3061,3062,3063,3064,3065,3066,3067,3068,3069,3070,3071,3072,3073,3074,3075,3076,3077,3078,3079,3080,3081,3082,3083,3084,3085,3086,3087,3088,3089,3090,3091,3092,3093,3094,3095,3096,3097,3098,3099,3100,3101,3102,3103,3104,3105,3106,3107,3108,3109,3110,3111,3112,3113,3114,3115,3116,3117,3118,3119,3120,3121,3122,3123,3124,3125,3126,3127,3128,3129,3130,3131,3132,3133,3134,3135,3136,3137,3138,3139,3140,3141,3142,3143,3144,3145,3146,3147,3148,3149,3150,3151,3152,3153,3154,3155,3156,3157,3158,3159,3160,3161,3162,3163,3164,3165,3166,3167,3168,3169,3170,3171,3172,3173,3174,3175,3176,3177,3178,3179,3180,3181,3182,3183,3184,3185,3186,3187,3188,3189,3190,3191,3192,3193,3194,3195,3196,3197,3198,3199,3200,3201,3202,3203,3204,3205,3206,3207,3208,3209,3210,3211,3212,3213,3214,3215,3216,3217,3218,3219,3220,3221,3222,3223,3224,3225,3226,3227,3228,3229,3230,3231,3232,3233,3234,3235,3236,3237,3238,3239,3240,3241,3242,3243,3244,3245,3246,3247,3248,3249,3250,3251,3252,3253,3254,3255,3256,3257,3258,3259,3260,3261,3262,3263,3264,3265,3266,3267,3268,3269,3270,3271,3272,3273,3274,3275,3276,3277,3278,3279,3280,3281,3282,3283,3284,3285,3286,3287,3288,3289,3290,3291,3292,3293,3294,3295,3296,3297,3298,3299,3300,3301,3302,3303,3304,3305,3306,3307,3308,3309,3310,3311,3312,3313,3314,3315,3316,3317,3318,3319,3320,3321,3322,3323,3324,3325,3326,3327,3328,3329,3330,3331,3332,3333,3334,3335,3336,3337,3338,3339,3340,3341,3342,3343,3344,3345,3346,3347,3348,3349,3350,3351,3352,3353,3354,3355,3356,3357,3358,3359,3360,3361,3362,3363,3364,3365,3366,3367,3368,3369,3370,3371,3372,3373,3374,3375,3376,3377,3378,3379,3380,3381,3382,3383,3384,3385,3386,3387,3388,3389,3390,3391,3392,3393,3394,3395,3396,3397,3398,3399,3400,3401,3402,3403,3404,3405,3406,3407,3408,3409,3410,3411,3412,3413,3414,3415,3416,3417,3418,3419,3420,3421,3422,3423,3424,3425,3426,3427,3428,3429,3430,3431,3432,3433,3434,3435,3436,3437,3438,3439,3440,3441,3442,3443,3444,3445,3446,3447,3448,3449,3450,3451,3452,3453,3454,3455,3456,3457,3458,3459,3460,3461,3462,3463,3464,3465,3466,3467,3468,3469,3470,3471,3472,3473,3474,3475,3476,3477,3478,3479,3480,3481,3482,3483,3484,3485,3486,3487,3488,3489,3490,3491,3492,3493,3494,3495,3496,3497,3498,3499,3500,3501,3502,3503,3504,3505,3506,3507,3508,3509,3510,3511,3512,3513,3514,3515,3516,3517,3518,3519,3520,3521,3522,3523,3524,3525,3526,3527,3528,3529,3530,3531,3532,3533,3534,3535,3536,3537,3538,3539,3540,3541,3542,3543,3544,3545,3546,3547,3548,3549,3550,3551,3552,3553,3554,3555,3556,3557,3558,3559,3560,3561,3562,3563,3564,3565,3566,3567,3568,3569,3570,3571,3572,3573,3574,3575,3576,3577,3578,3579,3580,3581,3582,3583,3584,3585,3586,3587,3588,3589,3590,3591,3592,3593,3594,3595,3596,3597,3598,3599,3600,3601,3602,3603,3604,3605,3606,3607,3608,3609,3610,3611,3612,3613,3614,3615,3616,3617,3618,3619,3620,3621,3622,3623,3624,3625,3626,3627,3628,3629,3630,3631,3632,3633,3634,3635,3636,3637,3638,3639,3640,3641,3642,3643,3644,3645,3646,3647,3648,3649,3650,3651,3652,3653,3654,3655,3656,3657,3658,3659,3660,3661,3662,3663,3664,3665,3666,3667,3668,3669,3670,3671,3672,3673,3674,3675,3676,3677,3678,3679,3680,3681,3682,3683,3684,3685,3686,3687,3688,3689,3690,3691,3692,3693,3694,3695,3696,3697,3698,3699,3700,3701,3702,3703,3704,3705,3706,3707,3708,3709,3710,3711,3712,3713,3714,3715,3716,3717,3718,3719,3720,3721,3722,3723,3724,3725,3726,3727,3728,3729,3730,3731,3732,3733,3734,3735,3736,3737,3738,3739,3740,3741,3742,3743,3744,3745,3746,3747,3748,3749,3750,3751,3752,3753,3754,3755,3756,3757,3758,3759,3760,3761,3762,3763,3764,3765,3766,3767,3768,3769,3770,3771,3772,3773,3774,3775,3776,3777,3778,3779,3780,3781,3782,3783,3784,3785,3786,3787,3788,3789,3790,3791,3792,3793,3794,3795,3796,3797,3798,3799,3800,3801,3802,3803,3804,3805,3806,3807,3808,3809,3810,3811,3812,3813,3814,3815,3816,3817,3818,3819,3820,3821,3822,3823,3824,3825,3826,3827,3828,3829,3830,3831,3832,3833,3834,3835,3836,3837,3838,3839,3840,3841,3842,3843,3844,3845,3846,3847,3848,3849,3850,3851,3852,3853,3854,3855,3856,3857,3858,3859,3860,3861,3862,3863,3864,3865,3866,3867,3868,3869,3870,3871,3872,3873,3874,3875,3876,3877,3878,3879,3880,3881,3882,3883,3884,3885,3886,3887,3888,3889,3890,3891,3892,3893,3894,3895,3896,3897,3898,3899,3900,3901,3902,3903,3904,3905,3906,3907,3908,3909,3910,3911,3912,3913,3914,3915,3916,3917,3918,3919,3920,3921,3922,3923,3924,3925,3926,3927,3928,3929,3930,3931,3932,3933,3934,3935,3936,3937,3938,3939,3940,3941,3942,3943,3944,3945,3946,3947,3948,3949,3950,3951,3952,3953,3954,3955,3956,3957,3958,3959,3960,3961,3962,3963,3964,3965,3966,3967,3968,3969,3970,3971,3972,3973,3974,3975,3976,3

Qy	421	GDPSVDIDNCNTCVTQTVDPSLDPTFIETTTVPQDAVSRSQRGRGTGRGMGLYRFVTPG	480
Db	1446		1505
		GDPSVDIDNCNTCVTQTVDPSLDPTFIETTTMPQDAVSRSQRGRGTSRGRGTYREVTPG	
Qy	481	ERPSGWFDSVLCBCEYDAGCAWVELTPAETSURLAYLNTPGLPVQODHLEFWESVFTGL	540
Db	1506		1565
		ERPSGWFDSVLCBCEYDAGCAWVELTPAETSURLAYLNTPGLPVQODHLEFWESVFTGL	
Qy	541	THIDAFLSQTQAGDNFFVLYAQATVCARAQAPPSWDMQMKCLIRLKPTLHGPTPLL	600
Db	1566		1625
		THIDAFLSQTQAGDNFFVLYAQATVCARAQAPPSWDMQMKCLTRLKPTLHGPTPLL	
Qy	601	YRLGAVONEVTTHTPIKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLITGCVVIVGR	660
Db	1626		1685
		YRLGAVONEVTLTHPIKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLITGCVVIVGR	
Qy	661	IILSGKPAITPREVLRYREFDEMEECASHLPYIEQGMQLAEQPKQKAILQLQATKQAEA	720
Db	1686		1745
		IILSGKPAVDPREVLRYQEFDEMEECASHLPYIEQGMQLAEQPKQKALGULQATKQAEA	
Qy	721	AAPVYESKWRTEAFWAKHWNFTISGIQYLAGLSTLPGNPAIASLAFTAITSPLTTQH	780
Db	1746		1805
		AAPVYESKWRTEAFWANDMWNFTISGIQYLAGLSTLPGNPAIASLAFTAITSPLTTQS	
Qy	781	TLFLNILGGWVAQALAPPSAASFPVGAGIAGAAGVSGIGLKVLDILAGVGAGVAGALVA	840
Db	1806		1865
		TLFLNILGGWVAQALAPPGAASAFPVGAGIAGAAGVSGIGLKVLDVMVAGYGAGVAGALVA	
Qy	841	FKWMSGMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGCBGAVQWMNRLIAPASR	900
Db	1866		1925
		FKWMSGMPSTEDLVNLLPAILSPGALVGVCAAILRRHVDPGEGAVQWMNRLIAPASR	
Qy	901	GNHVSPTHYPVESDAARVYQILSSLTITQLKRLHQMINEDCSTPCSGSWLRDVMWDIC	960
Db	1926		1985
		GNHVSPTHYPVESDAARVYQILSGLTITQLRRLHQMINEDCSTPCSGSWLRDVMWDIC	
Qy	961	TVLTDFTKWLQSKLLPRLPGVPFSCQRYKGVWRGDGIMQTTCPGQAOITGHVKGSGMR	1020
Db	1986		2045
		TVLADFTKLQSKLLPRLPGVPFSCQRYKGVWRGDGIMQTTCPGQAOITGHVKGSGMR	
Qy	1021	IVGPRTCSTNWHGTFPPINAYTTGPTCSPAPNYSRALWRVAEEYVEVTGVGFHYVTGM	1080
Db	2046		2105
		IVGPKTCSTNWHGTFPPINAYTTGCTCSPAPNYSRALWRVAEEYVEVRVGVGFHYVTGM	
Qy	1081	TTDNVCKPCQVPAPEFFEVDGVRLHRYAPACKPLLEEVTFVLVGLNQYLVGSLPCEPE	1140
Db	2106		2165
		TTDNVCKPCQVPAPEFFEVDGVRLHRYAPACKPLLEEVYFQVGLNQYVVGSQLPCEPE	
Qy	1141	PDVAVLTSMLTDPSHITAETAKRLARGSPPSLIASSASQLSAPSLKATCTTTRHSDPAD	1200
Db	2166		2225
		PDVAVLTSMLTDPSHITAETAKRLARGSPPSLIASSASQLSALSAACTTTRHTPPDAD	
Qy	1201	LIEANLLWROBMGNITRVSEKNKVILDSFEPLQAEDEREVSVPAEILRRSRKFFPRAM	1260
Db	2226		2285
		LIEANLLWROBMGNITRVSEKNKVILDSFDPLURAEDEREVSVPAEILRKSRKFFPAL	
Qy	1261	PIWARPNDYNPLLESWKDPDVPVPWVHGCPPLPPAKAPIPPPRKRTVWLSESTVSSALA	1320
Db	2286		2345
		PWARPNDYNPLLEPWKDPDVPVPVHGCPPLPVKAPAPIPPPRKRTVWLTESTVSSALA	
Qy	1321	ELATKTFGSSSESAVDSGTATASPDPQSDGDAGSDVESYSSMPLEGEPODPLSDGWS	1380
Db	2346		2405
		ELATKTFGSSSESAAGSGTATAPDPQSDGDAGSDVESCSSMPLEGEPODPLSDGWS	
Qy	1381	STVSEEBASEDVVCCSMSTYTGALITPCAABETKLPINALSNSLLRHNLVYATTSRSAS	1440
Db	2406		2465
		STVSEEDGEVIGCCSMSTYTGALITPCAABESKLPINALSNSLLRHNVMVYATTSRSAS	
Qy	1441	LROKQVTFDLRQVLDDHYRDVLKEMKAKASTVKAKLVSVEACKLTTPHARSKFGYGAK	1500
Db	2466		2525
		QROKQVTDRLQVLDDHYRDVLKEMKAKASTVKAKLVSVEACKLTTPHARSKFGYGAK	
Qy	1501	DVRNLSSKANVHNRSVMKDLLEDTEPTIDTTIMAKNEVFCVPQPKGGRKPARLIVFPDLG	1560

Db	2526	DVRNLGSKAHNHSRVWVKDLEETETPIDITIIWAKNEVFCVQPEKGGKPARLIVFFDLG	2589
Qy	1561	VRVCEKMALDYVYSTLTPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS	1620
Db	2586	VRVCEKMALDYVYSTLTPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS	2645
Qy	1621	TVTENDIRVEESIIYQCDDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLT	1680
Db	2646	TVTENDIRVEESIIYQCDDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLT	2705
Qy	1681	TSCGNTLTCVLKAAACRAAKLOCTMLVCGDDLWVCESAGTQEDASLRFAFTAMTRY	1740
Db	2706	TSCGNTLTCVLKAAACRAAKLOCTMLVCGDDLWVCESAGTQEDASLRFAFTAMTRY	2765
Qy	1741	SAPPDGPPEYDLELITSCSSNVSAHDASGRVYVYLTRDPTPLARAAWETARHTFPVN	1800
Db	2766	SAPPDGLPQPEYDELTITSCSSNVSAHDASGRVYVYLTRDPTPLARAAWETARHTFPVN	2825
Qy	1801	SWLGNITMYAPTLMARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL	1860
Db	2826	SWLGNITMYAPTLMARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL	2885
Qy	1861	HGLSAPSLHSYSGEINRVASCRLKGVPLRVWHRHARSVRALLSOGGAAATCGKYLIF	1920
Db	2886	HGLSAPSLHSYSGEINRVASCRLKGVPLRVWHRHARSVRALLSOGGAAATCGKYLIF	2945
Qy	1921	NWAVRTKRLKLTIPAAASQLDLSNFWAGYSGGDIYHLSRLRPRFWMCLLLSSVGVGIY	1980
Db	2946	NWAVRTKRLKLTIPAAASQLDLSNFWAGYSGGDIYHLSRLRPRFWMCLLLSSVGVGIY	3005
Qy	1981	LLPNR 1985	
Db	3006	LLPNR 3010	

RESULT 6

GNWV33

genome polyprotein - hepatitis C virus (strain HCV-1)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HCV-1)

C:Species: hepatitis C virus

C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004

C:Accession: A39166; PQ0403; PQ0404

R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos

Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991

A:Title: Genetic organization and diversity of the hepatitis C virus.

A:Reference number: A39166; MUID:91172826; PMID:1848704

A:Accession: A39166

A:Molecule type: mRNA

A:Residues: 1-3011 <CHO>

A:Cross-references: UNIPROT:p26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; P

B:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.

J. Gen. Virol. 73, 1131-1141, 1992

A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relation

A:Reference number: PQ0393; MUID:92268871; PMID:1316939

A:Accession: PQ0403

A:Molecule type: genomic RNA

A:Residues: 1577-1633 <CHA>

A:Cross-references: DDBJ:D10128

A:Experimental source: isolates E-b16

A:Accession: PQ0404

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1577-1633 <CH2>

A:Experimental source: isolates E-b17

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; non

F:1-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <BPM>

F:192-389/Product: major envelope protein E #status predicted <VEE>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>



Db 3005 IYLLPNR 3011  
|||||  
RESULT 7  
S40770  
genome polyprotein - hepatitis C virus  
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: hepatitis C virus  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: S40770; PC1285  
R;Okamoto, H.  
submitted to the EMBL Data Library, March 1992  
A;Reference number: S40770  
A;Accession: S40770  
A;Molecule type: genomic RNA  
A;Residues: 1-3011 <OKA>  
A;Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221586; R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Teuda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A;Reference number: PC1284; MUID:91013116; PMID:2170712  
A;Accession: PC1285  
A;Molecule type: genomic RNA  
A;Residues: 1-513 <OK2>  
A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512  
A;Experimental source: isolate HC-J1  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: Antp; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin  
F;2-115/Product: capsid protein C.#status predicted <CPC>  
F;116-191/Product: envelope protein M.#status predicted <EPM>  
F;192-389/Product: major envelope protein B.#status predicted <MBE>  
F;390-729/Product: nonstructural protein NS1.#status predicted <NS1>  
F;730-1006/Product: nonstructural protein NS2.#status predicted <NS2>  
F;1007-1615/Product: hepatitis C virus.#status predicted <NS3>  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEXH motif  
F;1616-1862/Product: nonstructural protein NS4a.#status predicted <N4A>  
F;1863-2013/Product: nonstructural protein NS4b.#status predicted <N4B>  
F;2014-3011/Product: nonstructural protein NS5.#status predicted <NS5>  
Query Match 89.8%; Score 9398.5; DB 1; Length 3011;  
Best Local Similarity 87.6%; Pred. No. 0;  
Matches 1740; Conservative 132; Mismatches 112; Indels 3; Gaps 3;  
QY 1 MAPTAYSQOTRGLGCIITSLTGRDRNOVEGEVQVSTQSFATCNGVCWTVYHGA 60  
Db 1026 LAPITAYQOTRGLGCIITSLTGRDRNOVEGEVQVSTQSFATCNGVCWTVYHGA 1085  
QY 61 GSKTLAGKGFITQMTYNVDQDLVGWQAPPGARSILPTCTCGSSDLYLVTRHADYIPVRRR 120  
Db 1086 GTRTIASPKGVIQWYTNVDQDLVGWQAPPGARSILPTCTCGSSDLYLVTRHADYIPVRRR 1145  
QY 121 GDSRGLSPRPVSYLKGSSGGPILCPGSHAVGIFRAVCTRGVAKAVDFVPVPSMETTM 180  
Db 1146 GDSRGLSPRPISYLGSSGGPILCPGSHAVGIFRAVCTRGVAKAVDFIPVPSLETTM 1205  
QY 181 RSPVFTNNSPPVQFQVAVHLHAPTGGSKSTKVPAAAYAGQYKVLVLPNSVAATLGF 240  
Db 1206 RSPVFTNNSPPVQFQVAVHLHAPTGGSKSTKVPAAAYAGQYKVLVLPNSVAATLGF 1265  
QY 241 AYMSKAHGIDNIRGTITTTGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 1266 AYMSKAHGIDNIRGTITTTGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 1325  
QY 301 ILGTGTVLDAQETAGARLVILATATPPGSVTVPHNPHEVALSSTGEIPFGKAIPIETI 360  
Db 1326 VLGTGTVLDAQETAGARLVILATATPPGSIIVPHANIEVALSSTGEIPFGKAIPIELAI 1385  
QY 361 KGRHLIFCHSKKKDELAALKLGLGNVAVYRGLDVSVIPTSGLDVVVVATDALMTGFT 420  
|||||

Db 1386 KGRHLIFCHSKKKDELAALKLGLGNVAVYRGLDVSVIPTSGLDVVVVATDALMTGFT 1445  
QY 421 GFDSVIDCNTCVTQVDFSLDPTFTTITTTVPODAVSRQRGRGTGRGMGIIVRFVTPG 480  
Db 1446 GFDSVIDCNTCVTQVDFSLDPTFTTITTTVPODAVSRQRGRGTGRGMGIIVRFVAPG 1505  
QY 481 ERSQGFDSVIDCEYDAGCAWVELTPEAETSVRLRAYLNTPLPVCODHLEFMSVETGL 540  
Db 1506 ERSQGFDSVIDCEYDAGCAWVELTPEAETSVRLRAYLNTPLPVCODHLEFMSVETGL 1565  
QY 541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWQMMKCLIRLKLPTLHGPTPL 600  
Db 1566 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWQMMKCLIRLKLPTLHGPTPL 1625  
QY 601 YRLGAVQNEVTTTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGWSVIVGR 660  
Db 1626 YRLGAVQNEVTTTHPIKTYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGWSVIVGR 1685  
QY 661 IILSGKPAITPDREVLYREDEMEECASHLPYIEQGMQLAEQPKAKGILLOTATKQAEA 720  
Db 1686 IILSGKPAITPDREVLYREDEMEECASHLPYIEQGMQLAEQPKAKGILLOTATKQAEA 1745  
QY 721 AAPVSESKWRTLBAFWAKHWNFIISGFIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db 1746 IAPVQTNWQKLEAFWAKHWNFIISGFIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 1805  
QY 781 TLLFNILGGVAAQALAPPASAASAVFGAGIAGAAGVSGTGLKVLVDIILAGYAGVAGALVA 840  
Db 1806 TLLFNILGGVAAQALAPPASAASAVFGAGIAGAAGVSGTGLKVLVDIILAGYAGVAGALVA 1865  
QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRIAPASR 900  
Db 1866 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRIAPASR 1925  
QY 901 GNVSPHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCTPCSGSLRDLVDWIMC 960  
Db 1926 GNVSPHYVPESDAAARVTQILSSLTITQLKRLHOWINEDCTPCSGSLRDLVDWIMC 1985  
QY 961 TVLTDFKTLQSKLLPRLPGVPFFSCQYKGVWVGIMQTTCCGQAOITGHVKNQSMR 1020  
Db 1986 EVLSDFKTLQSKLLPRLPGVPFFSCQYKGVWVGIMQTTCCGQAOITGHVKNQSMR 2045  
QY 1021 IVPRTCSNTHGTFPPINAVTTGCTPSAPNYSRALRWVAAAEVYEVTRVGDHYHVTGM 1080  
Db 2046 IVPRTCSNTHGTFPPINAVTTGCTPSAPNYSRALRWVAAAEVYEVTRVGDHYHVTGM 2105  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVLRLHRYAPACKPLLEEVTVLGLNQYLVGSLQCEPE 1140  
Db 2106 TTDNVKCPQVPAPEFFTEVDGVLRLHRYAPACKPLLEEVTVLGLNQYLVGSLQCEPE 2165  
QY 1141 PDVAVLTSMLTDPDSHITAEAKRLARGSPPSLAASSASQSLAPSLKATCTTRHDSDDAD 1200  
Db 2166 PDVAVLTSMLTDPDSHITAEAKRLARGSPPSLAASSASQSLAPSLKATCTTRHDSDDAD 2225  
QY 1201 LIEANLLWRQEMGNITRVSEKNNVILDSFELPQAEDEREVSVPAAEILRRSKFEPAM 1260  
Db 2226 LIEANLLWRQEMGNITRVSEKNNVILDSFELPQAEDEREVSVPAAEILRRSKFEPAM 2285  
QY 1261 PIWAPDYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRVTWZSESTVSSALA 1320  
Db 2286 PIWAPDYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRVTWZSESTVSSALA 2345  
QY 1321 ELATKTFGSSSSAVDSGTATASDPQSDDG - DAGSDVESYSSMPPLEGPGDPLSDSGS 1379  
Db 2346 ELATKTFGSSSSAVDSGTATASDPQSDDG - DAGSDVESYSSMPPLEGPGDPLSDSGS 2404  
QY 1380 WSTVSEEA - SEDVVCCSMSTWTGALITPCAAEETKLPINALSNLIRHNLVYATTSRS 1438  
Db 2405 WSTVSEEA - SEDVVCCSMSTWTGALITPCAAEETKLPINALSNLIRHNLVYATTSRS 2464  
QY 1439 ASLRQKVVTFDRLOVLDHVRDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYG 1498  
Db 2465 ASLRQKVVTFDRLOVLDHVRDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYG 2524





```
Db 1986 EVLSDFKTLKAKLMPQLGPIFFVSCQGRGVWRGDMHTRCHCAEITGHVKNQTM 2045
QY 1021 IVGPRTCSNTHGTFPFINATYTGCTPSPAPNYSRALRWAAAEYEVVTVRGDPHYVTGM 1080
Db 2046 IVGPRTCKNWSGTFFINATYTGCTPLPAPNYKFALRWYSAEYEVIRRVGDPHYVSGM 2105
QY 1081 TTDNVKCPQVPAPPEFTTEVDGVRHLHYAPACKPLLBREVTFLVGLNQYLVGSLPCEPE 1140
Db 2106 TTDNLKCPQTPSPFEFTTELDGVRHLHFAPCKPLLEEVSRFVGLHEYFVGSQLPCEPE 2165
QY 1141 PDVAVLTSMLTDPGSHITAETAKRELARGSPSLASSASQLSAPLSKATCTTTHSDPAD 1200
Db 2166 PDVAVLTSMLTDPGSHITAETAGRLARGSPSLMASSASQLSAPLSKATCTTANHSDAE 2225
QY 1201 LIEANLLWRQEMGNGNITRVESENKVVILDSFEPLQAEDEDEREVSPAELIRSRKFFRAM 1260
Db 2226 LIEANLLWRQEMGNGNITRVESENKVVILDSFDPPLVAEEDEREVSVPAAELIRKSRFPAL 2285
QY 1261 PIWARPDPNPLLESWKDPDYPVPVVGHCPLPPAKAPPIPPPRKRRTVVLSESTVSALA 1320
Db 2286 PWARPDYNPLLVETWKKPDYEPVHVHGCPLPPPRSPVPPPRKKRTVWLTESTLPTALA 2345
QY 1321 ELATKTFGSSSAVDGSTATASDPQSDDG-DAGSDVESYSSMPLEGGPDPLSDGS 1379
Db 2346 ELATKFGSSSTSGI-TGDNNTTSSEFAPGCPDSDVESYSSMPLEGGPDPLSDGS 2404
QY 1380 WSTVSEBA-SEDVVCCSMSYTWGTGALITPCAABETKLPIINALSNLSLRHNLVYATTSRS 1438
Db 2405 WSTVSSGADTEDVVCCSMSYTWGTGALVTPCAABEQKLPINALSNLSLRHNLVYATTSRS 2464
QY 1439 ASLRQKVTFDRLOVLDHVDLVKEMAKASTVAKALLSVEBEAKLTTPPHSARSKFGYG 1498
Db 2465 ACQRKKVTFDRLOVLDHVDLVKEMAKASTVAKALLSVEBEACSLAPPHSARSKFGYG 2524
QY 1499 AKDVRLNSSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 1558
Db 2525 AKDVRCARAKAVAHINSVWKKLLEDSVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 2584
QY 1559 LGVRVCEKMAIYDVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFYDTRCF 1618
Db 2585 LGVRVCEKMAIYDVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKTPMGLSYDTRCF 2644
QY 1619 DSTVTENDIRVEESYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQONCGYRRCRASGV 1678
Db 2645 DSTVTESDIRTEEAIIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCYRRCRASRV 2704
QY 1679 LTTSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMT 1738
Db 2705 LTTSCGNLTTRYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQDDASLRAFTAMT 2764
QY 1739 RYSAPGDPDPPEYDLELITSCSSNVSAHMDASGRVYVLTROPTTTLARAAMETARHTP 1798
Db 2765 RYSAPGDPDPPEYDLELITSCSSNVSAHMDAGKRVYVLTROPTTTLARAAMETARHTP 2824
QY 1799 VNSWLGNIIMVAPTLWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQII 1858
Db 2825 VNSWLGNIIMFAPTLWARMILMTHFFSVLTARDQLEALNCEIYGACYSIEPLDLPQII 2884
QY 1859 RLHGLSAFSLHSYSGEINRVASCLRLKLGVPPLRVWRHRSVRARLLSQGGAATCGKY 1918
Db 2885 RLHGLSAFSLHSYSGEINRVAAVCLRLKLGVPPLRAWHRWRSVRARLLRGGKAAICGY 2944
QY 1919 LFNWAVRTKLTPIPAASQDLSSWFFVAGYSGGDIYHLSRARPRFWMCLLSLVGVG 1978
Db 2945 LFNWAVRTKLTPIPAAGRLDLSGWFTAGYSGGDIYHVSVSHARPRFWFCLLLLAAGVG 3004
QY 1979 IYLLPNR 1985
Db 3005 IYLLPNR 3011
```

RESULT 9

```

JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9223423
A;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: UNIPROT:O39928; GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
A;Note: the translation of the nucleotide sequence is not complete in this paper
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;116-191/Product: capsid protein C #status predicted <CPC>
F;192-389/Product: major envelope protein E #status predicted <BPM>
F;384-408/Region: hypervariable #status predicted
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F;1008-1616/Product: hepatitis C virus NS2 #status predicted <NS2>
F;1231-1238/Region: nucleotide-binding motif A (P-loop)
F;1313-1318/Region: nucleotide-binding motif A (P-loop)
F;1317-1320/Region: DEXH motif
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4a>
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4b>
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F;2210-2245/Region: interferon sensitivity determining #status predicted
Query Match 80.4%; Score 8412.5; DB 1; Length 3014;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1545; Conservative 202; Mismatches 236; Indels 7; Gaps 3;
QY 1 MAPITAYSQTRGLGCIITSLGRDNRQVEGEVQVVSSTATQSFATCVNGVCMVTYHGA 60
Db 1027 LAPITAYAQTRGVGLGAILSLTRDKNEAGEVQFLSTATQTFGLGICINGVMVTLFHA 1086
QY 61 GSKTLAGPKPIQMTYNTVDOLVGOAPPGASLFTCTCGSSDLYLVTHADVIPRRR 120
Db 1087 GSKTLAGPKPVQMTYNTVDKDLVGFSPGKSLTCTCGSADLYLVTHADVIPARR 1146
QY 121 GDSRGSLLSRPVSYLKSGSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVBSMETTM 180
Db 1147 GDTFASLLSRPISYLYKSGSGGPI MCPGSHVGVFRAAVCTRGVAKALEFVVENLETTM 1206
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLFGF 240
Db 1207 RSPVFTDNTSPPAVPHEFQVGHLPHTGSGKSTKVPAAAYAAQGYKVLVLPNSVAATFGF 1266
QY 241 AYMSKAGHDPIRTGVTITTCAPTITGYKFLADGCGSGGAYDIIICDECHSTDTT 300
Db 1267 AYMSRAGVDPIRTGVTITTCAGITITGYKFFADGCGSGGAYDIIICDECHSQDATT 1326
QY 301 ILGIGTVLDDQAEATAGARLVVLTATPPGSVTVPHNPNEEVALSGTGPIFYGKAIPETI 360
Db 1327 ILGIGTVLDDQAEATAGARLVVLTATPPGSVTVTHNPNEEVALPSEGIPIFYGRAIPLVI 1386
QY 361 KGRGHLIFCHSKKKXKDELAALKSLGLNVAAYYRGDLVSVIPSGDVIVVATDALMTGFT 420
Db 1387 KGRGHLIFANQKKKAKETAKKMKPKGRKAVAYYRGDLVAVIPATGVDVVCSTDALMTGFT 1446
QY 421 GDPSVDICNTCTVOTQTFSLDPTFTTETTTTQDASVRSQRGRGMYRFRFTPG 480
Db 1447 GDPSVDICNSAVTQTQTFSLDPTFTTETTTTQDASVRSQRGRGMYRFRFTPG 1506
QY 481 ERPSGMFSDSVLCECYDAGCAWYELTAPETSRLRAYLNTPLPVCOHLEFWESEVFTGL 540
Db 1507 ERPSGIFSDSVLCECYDAGCAWYELTAPETSRLRAYLNTPLPVCOHLEFWESEVFTGL 1566
```



QY	541	THIDAHFLSOTKQAGDNFPYLVAOATVCARAQAPPEPWSQMKCLIRLXPTLHGTPPLL	600
DB			
DB	1567	TNIDAHMLSOAQOGGENFPYLVAOATVCRAKAPPPSWDTMWMCMICLXPTLTGTPTPLL	1626
QY	601	YRLGAVQNEVTTTHPIIKYINACMSADLEVVVTSVWLVGGVLAALAAAYCLITGSGVIVGR	660
DB			
DB	1627	YRLGAVQNEIITLTHPIIKYINACMSADLEVITSVWLVGGVVAALAAAYCLITGSGVAIVGR	1686
QY	661	IILSGKPAIIPDREVLRYEPDEMEECASHUPIYIEOGMQLAEQFKOKAIGLLOATKQAEA	720
DB			
DB	1687	IILSGRPAITFDREVILYQOPDEMEECASUPLYDEARAIAGQFKEKVLGLHIGTAGQKAEI	1746
QY	721	AAPVVESKWRITLEAPWAKHMNF:SGIYLAGUSTLPCNPAIASLMFTAFTASITSPLTQH	780
DB			
DB	1747	LKPAATSMWSKAEQFWAKHMNFVSGIYLAGUSTLPCNPAVATLMFTAFTASITSPLTQH	1806
QY	781	TLLENILGGWVAOALAPPSAASAFVGAGIAGAAVSG:GLGKVLVDIILAGYAGVAGALVA	840
DB			
DB	1807	TLLFNILGGWVASQAPPTAATAFVSGMAGAAVGN:IGLKVLIIDIIILAGYGTGVAGALVA	1866
QY	841	FKVMSGEMPTEDLVNLLPALISPGALVGVVCAAILRRHVGPGEVQWMMRLIAFASR	900
DB			
DB	1867	FKIMCGERPTEAELVNLLPSILCPGALVGVYCAVLRHLHIGPGEVQWMMRLIAFASR	1926
QY	901	GNHVSPHYVPESDAARVTOILSSLFI:TQLKRLHOWINEDCSTPCSGSMLRDVMDWIC	960
DB			
DB	1927	GNHGSPTHYVPETDASAKVTQLSSLTVTSLIKRLHTWI:GEDYSTPCDGTWLRALMDWVC	1986
QY	961	TVLTDFKTLQSLKLLPRLPGVPFPSCORGKYGWVRGDM:QOTTCPCGAQITGHVKNXGMR	1020
DB			
DB	1987	TALTDFKAWLKQKLLPQLPGVPFPSCORGKYGWVRGDM:KTCPCGATISGHVKNXTMR	2046
QY	1021	IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYGRALWRVAAAEYVEVTVRGDFHYVTGM	1080
DB			
DB	2047	IVGPKLCSNTWQGTFFINATTTGSPVAPAPNYKFA:LRVGAADYAEVRVGVHYITGV	2106
QY	1081	TTDNVKKCPQVPAPPEFFTEVDGVLRYAPACKP:LLREEVTVFVGLNQYLVGSQLPCEPE	1140
DB			
DB	2107	TQDMLKCPQVPSPPEFFTELDGVR:IHRFAPP:CNPLLEEVTVFVGLHSYVVGSQLPCEPE	2166
QY	1141	PDVAVLTSMLTDPSHITAEAKRRLARGSPPSLASSASOL:SPASLKATCTTTRHDSPAD	1200
DB			
DB	2167	PDVTVLSMLSDPAHITAEAKRRLNRGSPPSLANSSASOL:SPASLKATCTIOQHHPDAD	2226
QY	1201	LIEANLLWRQBMGNITRVSENVKVVILDSFEP:LOABEDEREVSVPABEILRRSRKFTPRAM	1260
DB			
DB	2227	LIKANLLWRQCMGNIITRVEAENKVEILDCFKPL:KEEDDREISVSADCFKGPAPFPAL	2286
QY	1261	PIWARPDPYNPPLLESWKDPDVPVPHGCLP:PPAKAPPIPPPRKR-TTVVLSSTVSSAL	1319
DB			
DB	2287	PWARP:GYPDPLLTWKRDPDYPQVWGC:PIPPAGPPFVPLPRKRKPMELSDSTVSQVM	2346
QY	1320	AELATKTF- - -GSSESASVDSGTATASDPQPSDDGAGSDVESYSMPLEGEPPDPL	1375
DB			
DB	2347	ADLADAREKVDYTP:IEQDQ:SA:GTS:SHQSDSPEKRDND:SDAASYSNMPLEGEPPDPL	2406
QY	1376	SDGWSVTVSBEASDVVCCMSYTWTGALITPCAABETKLPIN:ALNSLLRHNLYVATT	1435
DB			
DB	2407	SSGWSVTVSGS- -DNVWCCMSYTWTGALITPCAABEKLPI:NP:LSNTLLRHNLYVSTS	2464
QY	1436	SRSASLRQKVTFRLOVLDDHYEDVLKEMKAKAS:TVKALLSVEBACKLTPPHSARSKF	1495
DB			
DB	2465	SRSAGLRQKVTFRLOVLDDHYREVDEMKRLASKYKARLLPLEEACGLTPPHSARSKY	2524
QY	1496	GYGAKDVRNLSKSAVNHIRSYVMKOLLEDTETPIDTIT:MAKNEVFCVQPEKGGKPARLIV	1555
DB			
DB	2525	GYGAKEVRSLDKALKHIEGVQDLLDSDTDELFTT:IMAKNEVFAVEPSKGGKPARLIV	2584
QY	1556	FPDILGVRCEKMALYDVVSTLTPQAVMGSYGFQVSPQORVEFLVNAWAKKCPMGEFAYDT	1615
DB			
DB	2585	YPDILGVRCEKRALYDVAQKLP:ALMGPSYGFQVSPAQ:RVDLLKAWSKKIPMAFSYDT	2644

Qy	1616	RCFSTVTVENDIRVEBSIYQCCDLIAPEARQAIRSLTERLYIGGPLTNSKQNCQYRRCRA	16775
Db	2645	RCFSTTITEHDIMTEESIQSCDLQPEARVAIRSLTQRLYCGGPMWNSKQCGYRRCRA	2704
Qy	1676	SGVLTTSCGNLTCTYLKAAACRAAKLQDCTMLVCGDDLVAVICESAGTQDEASIRAFTE	1735
Db	2705	SGVFTTSGNMTCTYIKALASCAAKLQDCTLLVCGDDLVAVICESQTHDEASIRAFTE	2764
Qy	1736	AMTRYSAAPPDPPKPEYDLELITSCSSNVSAHDAASGRVYVYLTRDPTTTLAPAAAMETAR	1795
Db	2765	AMTRYSAAPPDPPPAYDLELVITSCSSNVSAVDASGNRIYYLTRDPQVPLAKAMETAK	2824
Qy	1796	HTPVNSWLGNIIMVAPTLWARMILMTHFFSLLIAEOLEKALCOQIVGACYSIEPDLDPQ	1855
Db	2825	HSPVNSWLGNIIMVAPTLWARIVLMTHTFFSVLOSQEQLTAFEMYGVSIVSTPLDLPA	2884
Qy	1856	IIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVPPPLRVWRHRSVRAIRLLSOGGAAATC	1915
Db	2885	IIQRLHGLSAFSLHSYSPSEINRVASCLRKLGVPPPLRAWRHRAVRRAKLIIOGGRAAIC	2944
Qy	1916	GKYLNFNAVTRTKLTLPIPAASQDLSSWTFVAGYSGGDIYHLSLRAPRPFMWCMLLLSV	1975
Db	2945	GIYLFNNAVTKRKLTLADADRLDLSWTFVAGGSDIYHMSRAPRNLLCLLLSV	3004
Qy	1976	GVGIYLLPNR	1985
Db	3005	GVGIYLLPAR	3014
RESULT 10			
QJ1303			
genome polyprotein - hepatitis C virus (isolate HC-J6)			
N:Contains: capsid protein C; envelope protein M; hepatitisvirin (EC 3.4.21.98) (no			
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5			
C:Species: hepatitis C virus			
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004			
C:Accession: QJ1303			
J.Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miy			
R. Gen. Virol. 72, 2697-2704, 1991			
A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated f			
A:Reference number: QJ1303; MUID:9204440; PMID:1658196			
A:Accession: QJ1303			
A:Molecule type: genomic RNA			
A:Residues: 1-3033 <OKA>			
A:Cross-references: UNIPROT:P26660; GB:D00944; NID:g221650; PIDN:BAA00792.1; PI			
A:Experimental source: isolate HC-J6 from a Japanese individual			
C:Superfamily: hepatitis C virus genome polyprotein			
C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinas			
F:2-115/Product: capsid protein C.#status predicted <CPC>			
F:116-191/Product: envelope protein M.#status predicted <EPM>			
F:192-389/Product: major envelope protein E.#status predicted <MEE>			
F:390-733/Product: nonstructural protein NS1.#status predicted <NS1>			
F:734-1010/Product: nonstructural protein NS2.#status predicted <NS2>			
F:1011-1619/Product: hepatitisvirin.#status predicted <NS3>			
F:1316-1321/Region: nucleotide-binding motif B			
F:1320-1323/Region: DEXH motif			
F:1620-1866/Product: nonstructural protein NS4a.#status predicted <N4a>			
F:1867-2017/Product: nonstructural protein NS4b.#status predicted <N4b>			
F:2018-3033/Product: nonstructural protein NS5.#status predicted <NS5>			
F:196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,125			
Query Match 76.8%; Score 8037.5; DB 1; Length 3033;			
Best Local Similarity 73.7%; Pred.No. 0;			
Matches 1485; Conservative 216; Mismatches 273; Indels 41; Gaps 6;			
Qy	1	MARITAYSQOTRGLLGCIITSLTGDRNRQVEGVQVWSTATQSFATCVNGVQWTVVHGA	60
Db	1030	LAPITAYAQOTRGLLGTVIVSMTCGRKTEQAGHIVLSTVTSQFLGTTISSVLVTVVHGA	1089
Qy	61	GSKTLAGKGPITOMYTNVDQDLVGWQAPPGASRLPTCTCGSSDLYLVITHADVIPVRRR	120
Db	1090	GKTLAGSRGPVTOMYSAGGDLVGMSPPGTKSLBECTCGVDLVLTATNADVIPARRR	1149

```

Db      2230  MVDANLF-----MGGDVTRIESESKVWVLDSDPWEERSLEPSIPSEYMLPKRFFPAL 2285
Qy      1261  PIWARDYNPLLESWKDPDYPVPHVHCCLPAPAKAPDIPPPRKRRTVWLSSESTVSALA 1320
Db      2286  PANARDYNPLVESWKRPDIQATVAGCALPPKKTPTPPRRRTVGLSESSIALDQ 2345
Qy      1321  ELATKTFG-----SSESAVDSGTATAGPDQPSDDGDAGSDVESYSSMPLEGE 1370
Db      2346  QLAIKFGQPPPGSDSLSTGADAADSGRT-PPDEL-----ALSTGSISSMPLEGE 2399
Qy      1371  GDDEL-----SDGSNSTVSEASESDVVCSSMSYTWGALITCA 1410
Db      2400  GDPLEPEQVELQPPPOGGVVTGSGSGSWSTCSEE-DDSVVCSSMSYTWGALITCSP 2458
Qy      1411  EETKLPINALNSLLRHNLVYATTSASLRQKKVTFDLQLVLDHYRDLVKEMKAKAS 1470
Db      2459  EEEKLPINPLNSLLRYHNKYVCTTSKASLRKAKVTFDRMQALDAHVDVSKDKLAAS 2518
Qy      1471  TVKAKLSVEEACKLTPPHSARSKFGYKADVRNLSKAVNHRSVWKDLEDETETPIDT 1530
Db      2519  KVTARLLTLEEAQCLTPPHSARSKYFGAKEVRSLSGRAVNHKSVWKDLEDETQTPIT 2578
Qy      1531  TIWAKNEVFCQPEKGRKPARLIVFPDLGVRCVKMALVDVVTSLPQAVMGSSYGFQYS 1590
Db      2579  TIWAKNEVFCVDTKGGKKAARLIVPDLGVRCVKMALVDITQKLPQAVMGASYGFQYS 2638
Qy      1591  PGRVPEFLVNAWKAKKCPMGCFAYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQATRSL 1650
Db      2639  PAQVPEFLKAWAEKDDPMGFSVDTRCFDSTVTVERDIRTBESIYRACSLPEEAHTAHS 1698
Qy      1651  TERLYTGGPLTNSKGQNGCYRRCRASGLVTTSCGNTLTCTYLKAAAAACRAAKLQDCTMLVC 1710
Db      2699  TERLYVGGPMFNSKGQTCGYRRCRASGLVTTSMGNTITCTYKALAAKAAAGIIAPTMLVC 2758
Qy      1711  GDDLIVICESAGTQEDBASIRAFTEAMTRYSAAPGDPKPEYDELELTSCSSNVSVAHDA 1770
Db      2759  GDDLIVISESQGTEEDERNRAFTTEAMTRYSAAPGDPKPEYDELELTSCSSNVSVALGP 2818
Qy      1771  SGRKVVYLTDRPTTPLARAWEATRHPTPVNSWLGNIIMYAPTLMWMLTMHFFSIILAQ 1830
Db      2819  QGRRYVYLTDRPTTPIARAWEATVRHSPVNSWLGNIIOYAPTIIMWMLTMHFFSILMQ 2878
Qy      1831  EQLEKALDCQTYGACYSEPLDLPQIQRHLGLSAFSLHSYSPGEINRVASCLRKLGVPP 1890
Db      2879  DTLQNLNLFEMYGAVYSVPLDLPALIERLHGLDAFSLHTYTPHELTRVASAKKLGA 2938
Qy      1891  LRWRHRAVSVRAPLLSQGGAATCGKYLFWAVRTKLTPTPAASQLDLSHFVAGYS 1950
Db      2939  LRWAKSRARAVRASLI SRGGRAAVCGRYLFWNAVKTKLTPLEAPRLDLSWFTVGAG 2998
Qy      1951  GGDYIHSLSRARPRFWMKCLLLLSVGVGIVLLPNR 1985
Db      2999  GGDYIHSVSARPELLELLGLLLLVGVLFLPAR 3033

RESULT 11
GNWVJ8
genome polyprotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepatitisvirin (EC 3.4.21.98) (m
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40250; PQ0397; F00559
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fu
Virology 188, 331-341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homolog
A:Reference number: A40250; MUID:92230232; PMID:1314459
A:Accession: A40250
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OK>
A:Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:G221608; PIDN:BA
R:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.;

```



Db 2643 VDFLLKAWGSKDPMPGFSYDTRCFDSTVTERDIRTEESIIYQACSLPQEARVTHLSLTERL 2702

QY 1655 YIGPLTNSGONGCYRRCRASGVLTTSCGNTLTCYIKAAACRAKLOCTMLVCGDDL 1714

Db 2703 YVGEMTNSGQSGCYRRCRASGVFTTSMGNTMTCYIKALAAACAAAGIVPVMVLCGGDDL 2762

QY 1715 VICESAGTODEASLRAFTEAMTRYSAPOCDPKPEYDLELITSCSSNVVAHDASGR 1774

Db 2763 VVISEOGNEEDERLRAFTEAMTRYSAPOCDLPRPEYDLELITSCSSNVVALDSRGR 2822

QY 1775 VYILTRDPTTLARAAWETARTHPVNSWLGNIIMYAPTLWARMILMTHFTSILLAQOOLE 1834

Db 2823 RYFLTRDPTTLTRAAWETVRHSPVNSWLGNIIOYAPTIVRWVIMTHFTSILLAQOOLN 2882

QY 1835 KALDCQIYGACYSIEPLDLPOIQRHLGSAFSLHSYSGEINRVASCLRLKLGVPPLRVW 1894

Db 2883 QNLNFEMYGAVSVNPLDLPAIIRLHGLBAFSLHTYSPELSRVAATLRLKLGAPPLRAW 2942

QY 1895 RHPARSVRARLLSOGGAAATCGKYLEFNWAVTKLTPIPAASOLDLSSWFAVYSGGDI 1954

Db 2943 KSRARVRASLIQAARAAICGRYLFNVAVKTLKLPPLPEASRLDLSGWFVAGGGDI 3002

QY 1955 YHLSLRARPWFWMCLLLLSGVGYIYLLPNR 1985

Db 3003 YHVSHPARPLLLCLLLLSGVGIFLLPAR 3033

RESULT 12

PS0326

polyprotein - hepatitis C virus (isolate Fla) (fragments)

C:Species: hepatitis C virus

C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004

C:Accession: PS0326

R:Li, J.S.; Tong, S.P.; Vitvitski, L.; Lepot, D.; Trepo, C.

Gene 105, 167-172, 1991

A:Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype

A:Reference number: PS0326; MUID:92039028; PMID:1718820

A:Accession: PS0326

A:Molecule type: genomic RNA

A:Residues: 1-492 <lit>

A:Cross-references: UNIPROT:Q91FES; UNIPROT:Q36579; UNIPROT:Q36610; UNIPROT:Q03463; UNIPROT:M60220

A:Note: this sequence corresponds to nonstructural protein NS3 region

A:Translation of the nucleotide sequence is not complete

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: polyprotein

Query Match 22.8%; Score 2383; DB 2; Length 492;

Best Local Similarity 90.0%; Pred. No. 7e-138;

Matches 443; Conservative 28; Mismatches 21; Indels 0; Gaps 0;

QY 425 SVIDCNCVTQTVDPSLDPTFTTTPQDAVSRSQRRGTGRGMGIYRFVTPGERPS 484

Db 1 SVIDCNCVTQTVDPSLDPTFTTTPQDAVSRTQRRGTGRKGPIYRFVAPGERPS 60

QY 485 GMPDSSVLCYDAGCAWYELTPAETSVRILAYINTPGLVPCQDHLFPWESVTGLTHID 544

Db 61 GMPDSSVLCYDAGCAWYELTPAETVRLAYINTPGLVPCQDHLFPWESVTGLTHID 120

QY 545 AHFLSOTKQAGDNFPYLVAYQATVCARAQAPPPSWDQMKLIRKPLTHGTPPLLYRLG 604

Db 121 AHFLSOTKQSGENLPYLVAQATVCARAQAPPPSWDQMKLIRKPLTHGTPPLLYRLG 180

QY 605 AVQNEVTHTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTSGWIVGRILS 664

Db 181 AVQNEITLTHPTIKYIMTCMSADLEVVTSTWLVGGVLAALAAAYCLTGTSGWIVGRILS 240

QY 665 GKPAIPDRVLRPEFDEMECSHLPYIEGQHLAQFOKQALGILLQATKQAEAAAPV 724

Db 241 GKPAIPDRVLRPEFDEMECSHLPYIEGQHLAQFOKQALGILLQATKQAEAAITPA 300

QY 725 VESKWRITLFAFWAKHMWNFIISGQIYLAGLSTLPCNPAIASIMAFATISPLTQHTLLF 784

Db 301 VQTNWQBLETFWAKHMWNFIISGQIYLAGLSTLPCNPAIASIMAFATISPLTQHTLLF 360

QY 785 NILGGWVAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDIILAGYAGVAGALVAFKVM 844

Db 361 NILGGWVAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDIILAGYAGVAGALVAFKVM 420

QY 845 SGEMPSTEDLWLLPAILSPALVGVVCAAILRRHVPGEAGVQWNNRLIAFASRGNHV 904

Db 421 SGEPSTEDLWLLPAILSPALVGVVCAAILRRHVPGEAGVQWNNRLIAFASRGNHV 480

QY 905 SPTHYVPESDAA 916

Db 481 SPTHYVPESDAA 492

RESULT 13

PC2219

polypeptide - hepatitis C virus (type 5a) (fragments)

N:Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A protein

C:Species: hepatitis C virus

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: PC2219

R:Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.

Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994

A:Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the hepatitis C virus genome

A:Reference number: PC2219; MUID:94338342; PMID:7520237

A:Accession: PC2219

A:Molecule type: mRNA

A:Residues: 1-876 <STU>

A:Cross-references: UNIPROT:Q81242; GB:L29577; GB:L29578; GB:L29579

A:Experimental source: serum

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: glycoprotein

F:1-191/Product: core #status predicted <COE>

F:68-78/Region: variable

F:192-247/Product: E1 (carboxyl end) #status predicted <BRE>

F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>

F:248-338/Region: E2

F:339-411/Region: NS1 (amino end)

F:412-783/Product: NS3 #status predicted <NSR>

F:784-837/Product: NS4A #status predicted <NSB>

F:838-876/Product: NS4B #status predicted <NSB>

F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.0%; Score 2098; DB 2; Length 876;

Best Local Similarity 83.0%; Pred. No. 4.6e-120;

Matches 385; Conservative 39; Mismatches 40; Indels 0; Gaps 0;

QY 260 ITTGAPITSTYCKFLADGCGSGAYDIIICDCHSTDSITILGIGTIVLDOAETAGARLV 319

Db 411 ITTGASITSTYCKFLADGCGSGAYDIIICDCHSTDSITILGIGTIVLDOAETAGARLV 470

QY 320 VLATATPPGSVTPHNIEVALSSGEIPFYKAIPIETIKGRHLIFCHSKKCCDELA 379

Db 471 VLATATPPGSVTPHNIEVALPQGEVFPYGRALPLAFIKGRHLIFCHSKKCCDELA 530

QY 380 AKLSGLNNAVAYRGLDVSIVPTSGDIVVATDALMTGTGDFSDVIDCNTCVTQTVDF 439

Db 531 KQLTSLGVNAVAYRGLDVAIVTAGDVVVCSTDALMTGTGDFSDVIDCNSAVTQTVDF 590

QY 440 SLDPPTIETITTPQDAVSRSQRRGTGRGMGIYRFVTPGERPSGFMFSSVLCYCDAG 499

Db 591 SLDPPTIETITTPQDAVSRSQRRGTGRGMGIYRFVTPGERPSGFMFSSVLCYCDAG 650

QY 500 CAWVELTPAETSVRILAYINTPGLVPCQDHLFPWESVTGLTHIDAHFLSOTKQAGDNRP 559

Db 651 CAWVELTPAETSVRILAYINTPGLVPCQDHLFPWESVTGLTHIDAHFLSOTKQAGDNRP 710

QY 560 YLVAYQATVCARAQAPPPSWDQMKLIRKPLTHGTPPLLYRLGAVQNEVTHTHPTIKY 619

Db 711 YLVAYQATVCARAQAPPPSWDQMKLIRKPLTHGTPPLLYRLGAVQNEVTHTHPTIKY 770

QY 620 IMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTSGWIVGRILSKPAIIPDRVLYRE 679





QY 1331 -----ESSAVDSGTATASPDQDDGD-----AGSDVESYSSMPLEGEPPDLDSDGSM 1380  
 Db 2337 RILRVQGCCDRSTLUKAFPLE-LKVGEVLFVCGYDPEGHR-----VTDGDDDMN-----2387  
 QY 1381 STVSBEASEDDVVC-----SMSTWTGA-----LITPCAAEETKLPINALSNSLLR 1426  
 Db 2388 TFIADTLGDIILSCETREBAQXSYSYWSGAPLGCGRVVP-----PITRPIGTHLT 2439  
 QY 1427 HHNL-VYATTSASLRQKKVTFDRLO-VLDHYRDVLKEMKAKASTVAK-----LLSVE 1480  
 Db 2440 HDTTKVYVTDPPRAERAERAKVTIWRQRYDEHYASVGEALKKAATSPGWTYMAIS 2499  
 QY 1481 EACKLTPPHSARSKFGYAKOVNRLSSKAVNHRSVWKDLLEDTEPIDTTIMAKNEVFC 1540  
 Db 2500 KVRK-----RAAGYGSKVTVATLETGWPHVRTIMDKIRCHEEVPF--TFVTKREVEFF 2550  
 QY 1541 VOPEKGRKPARLIYPPDLGVRVCERKALYDVVSTLPQAVMGSSYGFQYSPGQORVEFLVN 1600  
 Db 2551 ---SKTRKPPRFIVVPPDLFRIAEKMLGD-PGIVAKAILGPAYLFQVTPNQRVKLLVS 2606  
 QY 1601 AWKAKCPMGFAIDRCFSTVTENDIRVEBSIYQCCDLAPARQAIRSLTERLYIGGPL 1660  
 Db 2607 TWESKTHPAATVDTATCFDSSIDERDMEVEAEIFAASDQPELVKALC----RYAEGEM 2662  
 QY 1661 TNSKGQMGYRRCRASGVLTTCGNTLTCLYKAAAAACRAAKLQDCTMLVCGDDLTVICES 1720  
 Db 2663 VTPDGVPIGFRKCRSGVLTSSANSITCYIKVKAACAKVGLKDPSPFFIAGDDCLIIYED 2722  
 QY 1721 AGTQDEASLRAFTAMTRYAPPDPPKPEYDLELITS-CSSNVSVAHDAKRVYIYL 1779  
 Db 2723 DGADPCE-RURL---ALGNY---GYRCEPSKHASLDTAECCSAYLAECTAGGSRRWLS 2774  
 QY 1780 RPTTFLARAAMETARHTPVNSWLGNIIMVAPTLWARMILMTHFFSILLAQE---OLEKA 1836  
 Db 2775 TDMRKELARAAAEYS--DFVGSALGTILMPWHPVIVRYVLIPIHV--LIMAFRGGGTPDEL 2830  
 QY 1837 LDCQIYGACYSIEPLDLPQIQLHGLSAPSLHSYSPGEINRVASCLKLGVPPLRVWRH 1896  
 Db 2831 VMCQVQNTYSPFLKLPVLVSLHGPWCLQVTTDSTKTEMEAGNALRDLGKXLSWHRR 2890  
 QY 1897 RARSVRARLLSQO---GRAATCGKYLEFNWAVTKLKLTPIPASQLDLSWVYAGSGD 1953  
 Db 2891 RAGNVTRLLRGKAWGHLARC---LLWHPGLKEHPPPVIALPNFRL----VTPVEHHE 2942  
 QY 1954 IYHLSRAPRPRFMWCL 1970  
 Db 2943 EVLISLKERPPWVKWLL 2959

Search completed: December 8, 2004, 12:33:12  
 Job time : 70 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:14:57 ; Search time 272 Seconds  
(without alignments)  
4198.966 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQQTRELGLGCIIT.....FMWCLLLSVGVGIIYLLPNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10462	100.0	3010	2	Q9WMX2
2	10365	99.0	3010	2	Q9QP06
3	10245	97.9	3010	2	Q9J3H2
4	10225	97.7	3008	2	Q9J3F4
5	10225	97.7	3010	2	Q8V638
6	10218	97.6	3010	2	Q93016
7	10218	97.6	3010	2	Q9DTE7
8	10206	97.5	1984	2	Q7T4V8
9	10199	97.5	3010	2	Q90194
10	10194	97.4	3010	2	Q99AU2
11	10187	97.3	3010	2	Q9DTE9
12	10183	97.3	3010	2	Q9QIY8
13	10182	97.3	3010	2	Q90193
14	10177	97.2	3010	2	Q9J3H3
15	10177	97.2	3010	2	Q9QIX6
16	10177	97.2	3010	2	Q9DIF0
17	10176	97.2	3010	2	Q9J3G7
18	10176	97.2	3012	2	Q9QIX5
19	10172	97.2	3010	2	Q9QIX5
20	10172	97.2	3010	2	Q9QIY4
21	10172	97.2	3010	2	Q9DTE6
22	10170	97.2	3010	1	FOLG_HCVBK
23	10169	97.2	3010	2	Q9QIY3
24	10165	97.1	3010	2	Q9QIY7
25	10165	97.1	3010	2	Q9DTE6
26	10161	97.1	3010	2	Q9J3H9
27	10160	97.1	3010	2	Q9DTE2
28	10159	97.1	3010	1	FOLG_HCVJT
29	10157	97.1	3010	2	Q91AU0
30	10151	97.0	3013	2	Q9QIY0
31	10151	97.0	3014	2	Q6GYR8

32	10150	97.0	3010	2	Q9WIK8	Q9wik8 hepatitis c
33	10149	97.0	3010	2	Q9J3I1	Q9j3i1 hepatitis c
34	10148	97.0	3010	2	Q6GYR9	Q6gyr9 hepatitis c
35	10147	97.0	3010	2	Q807P3	Q807p3 hepatitis c
36	10147	97.0	3010	2	Q9J3G6	Q9j3g6 hepatitis c
37	10147	97.0	3010	2	Q9J3I0	Q9j3i0 hepatitis c
38	10145	96.9	3010	2	Q9J3G9	Q9j3g9 hepatitis c
39	10144	96.9	3013	2	Q9QIX9	Q9qix9 hepatitis c
40	10141	96.9	3010	2	Q9J3H7	Q9j3h7 hepatitis c
41	10141	96.9	3010	2	Q9QIY1	Q9qiyl hepatitis c
42	10136	96.9	3010	2	Q9DTE4	Q9dte4 hepatitis c
43	10132	96.8	3010	2	Q9QP61	Q9qp61 hepatitis c
44	10131	96.8	3010	2	Q9QIX2	Q9qix2 hepatitis c
45	10130	96.8	3010	2	Q09796	Q09796 hepatitis c

## ALIGNMENTS

RESULT 1  
Q9WMX2  
ID AC Q9WMX2; PRELIMINARY; PRT; 3010 AA.  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus type 1b.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 1.  
OX NCBI\_TaxID=31647;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RB;  
RX MEDLINE=99322193; PubMed=10390360;  
RA Lohmann V, Koerner F., Koch J.O., Herian U., Theilmann L.,  
RA Bartenschlager R.;  
RT "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line.";  
RL Science 285:110-113(1999).  
DR PIR; A61196; A61196.  
DR PIR; P00246; P00246.  
DR PIR; PS0329; PS0329.  
DR HSP; P26663; IUXP.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:virial envelope; IEA.  
DR GO; GO:0019031; C:virial envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0019079; P:virial genome replication; IEA.  
DR GO; GO:0019087; P:virial transformation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR InterPro; IPR000745; HCV\_NS4a.  
DR InterPro; IPR001490; HCV\_NS4b.  
DR InterPro; IPR002868; HCV\_NS5a.  
DR InterPro; IPR002166; HCV RdRp.  
DR InterPro; IPR001650; Helicase C.  
DR InterPro; IPR004109; Peptidase S29.  
DR InterPro; IPR009003; Pept Ser Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR	InterPro; IPR007094; RNA_pol_PSVir.	
DR	Pfam; PF01543; HCV capsid; 1.	
DR	Pfam; PF01542; HCV core; 1.	
DR	Pfam; PF01539; HCV env; 1.	
DR	Pfam; PF01560; HCV NS1; 1.	
DR	Pfam; PF01538; HCV NS2; 1.	
DR	Pfam; PF02907; HCV NS3; 1.	
DR	Pfam; PF01006; HCV NS4a; 1.	
DR	Pfam; PF01001; HCV NS4b; 1.	
DR	Pfam; PF01506; HCV NS5a; 1.	
DR	Pfam; PF00271; Helicase C; 1.	
DR	Pfam; PF00998; Viral_RdRp; 1.	
DR	SMART; SM00487; DEXDC; 1.	
DR	PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.	
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;	
KW	Polyprotein; Transmembrane.	
FT	CHAIN 810 1026 NS2 proteinase.	
FT	CHAIN 1027 1657 NS3 proteinase/helicase.	
FT	CHAIN 1658 1711 NS3/4A proteinase cofactor.	
FT	CHAIN 1 191 core protein.	
FT	CHAIN 1712 1972 NS4B protein.	
FT	CHAIN 1973 2419 NS5A phosphoprotein.	
FT	CHAIN 2420 3010 NS5B RNA dependant RNA polymerase.	
FT	CHAIN 192 383 Glycoprotein E1.	
FT	CHAIN 384 746 Glycoprotein E2.	
FT	CHAIN 747 809 p7 protein.	
SQ	SEQUENCE 3010 AA; 326903 MW; 37B3448DAFA9A10B CRC64;	
Query Match 100.0%; Score 10462; DB 2; Length 3010;		
Best Local Similarity 99.9%; Pred. No. 0;		
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MAPITAYSQOTRGLLCITSLTGRDRNQVEGVQVWVSTATQSFLLATCNGVCWTVYHGA 60	
DB	1026 LAPITAYSQOTRGLLCITSLTGRDRNQVEGVQVWVSTATQSFLLATCNGVCWTVYHGA 1085	
QY	61 GSKTLAGPKPITOMYTNVDQVLGVQAPPGARSLTPCTCGSSDLVLTVRHADVIPVRR 120	
DB	1086 GSKTLAGPKPITOMYTNVDQVLGVQAPPGARSLTPCTCGSSDLVLTVRHADVIPVRR 1145	
QY	121 GDSRGILLSRPVSYLKSGSGGLLCPSGHVGIFRAAVCTRGVAKAVDFPVVSMETTM 180	
DB	1146 GDSRGILLSRPVSYLKSGSGGLLCPSGHVGIFRAAVCTRGVAKAVDFPVVSMETTM 1205	
QY	181 RSPVFTDINSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGYKVLVLPNSVAATLFGF 240	
DB	1206 RSPVFTDINSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGYKVLVLPNSVAATLFGF 1265	
QY	241 AYMSKAGIDPNRTGVRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT 300	
DB	1266 AYMSKAGIDPNRTGVRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT 1325	
QY	301 ILGIGTVLDAQETAGARLVVLATATPGSVTVVPHNIEEVALSGTGEIPFYKAIPIETI 360	
DB	1326 ILGIGTVLDAQETAGARLVVLATATPGSVTVVPHNIEEVALSGTGEIPFYKAIPIETI 1385	
QY	361 KGRHLIFCHSKKKDBLAKLSGLGNVAVYRGLDVSVIPTSGDVIVVATDALMTGFT 420	
DB	1386 KGRHLIFCHSKKKDBLAKLSGLGNVAVYRGLDVSVIPTSGDVIVVATDALMTGFT 1445	
QY	421 GDPDSVTDNCNTQTVDFSLDPTFTTETTVQDAVRSQRGRTGRGMGIYRFVTPG 480	
DB	1446 GDPDSVTDNCNTQTVDFSLDPTFTTETTVQDAVRSQRGRTGRGMGIYRFVTPG 1505	
QY	481 ERPSGMFDSVLCBCEYDAGCAWYBELTPEATSVRLRAYLNTPGLPVQCDDHLEFWESVFTGL 540	
DB	1506 ERPSGMFDSVLCBCEYDAGCAWYBELTPEATSVRLRAYLNTPGLPVQCDDHLEFWESVFTGL 1565	
QY	541 THIDAHFLSQTQAGDNFPVLVAYQATVCARAQAPPSWDMKCLRLKLTLLHGPTLL 600	
DB	1566 THIDAHFLSQTQAGDNFPVLVAYQATVCARAQAPPSWDMKCLRLKLTLLHGPTLL 1625	
QY	601 YRLGAVQNEVTTTHITKIYIMACMSADLEVVTTSTWLVGGVLAALAAAYCLTTGSSWIVGR 660	
DB	1681 YRLGAVQNEVTTTHITKIYIMACMSADLEVVTTSTWLVGGVLAALAAAYCLTTGSSWIVGR 1740	
DB	1626 YRLGAVQNEVTTTHITKIYIMACMSADLEVVTTSTWLVGGVLAALAAAYCLTTGSSWIVGR 1685	
QY	661 IILSGKPAIIPDREVLYREFDEMERCASHLPVIEGQMQLAEQFKOKAIGLLQATKQABA 720	
DB	1686 IILSGKPAIIPDREVLYREFDEMERCASHLPVIEGQMQLAEQFKOKAIGLLQATKQABA 1745	
QY	721 AAPVYESKWRITLEAPWAKHMNFISGIQYLAGLSLTPGNPAIASLMAFTASITSPITTOH 780	
DB	1746 AAPVYESKWRITLEAPWAKHMNFISGIQYLAGLSLTPGNPAIASLMAFTASITSPITTOH 1805	
QY	781 TLLFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVALVA 840	
DB	1806 TLLFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVALVA 1865	
QY	841 FKVMGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVMMNRLIAFASR 900	
DB	1866 FKVMGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVMMNRLIAFASR 1925	
QY	901 GNHVSPTHVPSDAAARVTQILSSLITITQLLKRHLQWINEDCSTPCSGSWLRDVMWDMIC 960	
DB	1926 GNHVSPTHVPSDAAARVTQILSSLITITQLLKRHLQWINEDCSTPCSGSWLRDVMWDMIC 1985	
QY	961 TVLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDGIMQTTCPGCAQITGHVKNKSWR 1020	
DB	1986 TVLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDGIMQTTCPGCAQITGHVKNKSWR 2045	
QY	1021 IVGPRTCSNTHGTPPINAYTTGPTSPAPNYSRALWVAEEYVEVTVGDFHVTGM 1080	
DB	2046 IVGPRTCSNTHGTPPINAYTTGPTSPAPNYSRALWVAEEYVEVTVGDFHVTGM 2105	
QY	1081 TTDNVKCPQVPAPEEFTVDGVRLHRYAPACKPILRREEVTVLGLNQVLVGSQLPCEPE 1140	
DB	2106 TTDNVKCPQVPAPEEFTVDGVRLHRYAPACKPILRREEVTVLGLNQVLVGSQLPCEPE 2165	
QY	1141 PDVAVLTSLMTDFSHITAEAKRRLARGSPPLASSASSASQLGAPSILKATCTTHDSDPAD 1200	
DB	2166 PDVAVLTSLMTDFSHITAEAKRRLARGSPPLASSASSASQLGAPSILKATCTTHDSDPAD 2225	
QY	1201 LIEANLLWQBMGNITRVESENKVILDSPEPQAEDEDEREVSVAELIRRSRKPFRAM 1260	
DB	2226 LIEANLLWQBMGNITRVESENKVILDSPEPQAEDEDEREVSVAELIRRSRKPFRAM 2285	
QY	1261 PIWAREDYNPPLLESWKDPDYVPVHVHGPPLPAKAPPIPPPRRKKTIVLSESTVSALA 1320	
DB	2286 PIWAREDYNPPLLESWKDPDYVPVHVHGPPLPAKAPPIPPPRRKKTIVLSESTVSALA 2345	
QY	1321 ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPDPLSDGSW 1380	
DB	2346 ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPDPLSDGSW 2405	
QY	1381 STVSEASEDVCSSSYTWTGALITPCAABETKLPINALNSLLRHNILVYATTSSRSAS 1440	
DB	2406 STVSEASEDVCSSSYTWTGALITPCAABETKLPINALNSLLRHNILVYATTSSRSAS 2465	
QY	1441 LRQKVTFRQLVDLDDHYRDVLKEMKAKASTVAKALLSVEEACKLTPHPSARSKFYGAK 1500	
DB	2466 LRQKVTFRQLVDLDDHYRDVLKEMKAKASTVAKALLSVEEACKLTPHPSARSKFYGAK 2525	
QY	1501 DVNRLSSKAVNHIRSVNKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560	
DB	2526 DVNRLSSKAVNHIRSVNKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 2585	
QY	1561 VRVCEKXALYDVYSTLFOAVMGSSYGFQYSPGQRFVFLVNAWKAKKCPMGFADYDTRCFDS 1620	
DB	2586 VRVCEKXALYDVYSTLFOAVMGSSYGFQYSPGQRFVFLVNAWKAKKCPMGFADYDTRCFDS 2645	
QY	1621 TVTENDIRVEESYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYVRCASGVLT 1680	
DB	2646 TVTENDIRVEESYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYVRCASGVLT 2705	
QY	1681 TSCNTLTVCYKAAACRAAKLODCTMLVCGDDLVVICESAGTQEDASIRAFTEATRY 1740	

Db 2706 TSCGNTLTCLYLKAAACRAAKLQDCTMLVCGDVLVVICESAGTQDEASLRAFTEAMTRY 2765  
Qy 1741 SAPGDPKPKPYDLELITSCSNVSVVAHDAGSKRYVYITRPTTPPLAANAWEARHTPVN 1800  
Db 2766 SAPGDPKPKPYDLELITSCSNVSVVAHDAGSKRYVYITRPTTPPLAANAWEARHTPVN 2825  
Qy 1801 SWLGNIMYAPTLWARMILMTHFFSILLAOEKLKALDQCIYGACYSIEPLDLPQIQR 1860  
Db 2826 SWLGNIMYAPTLWARMILMTHFFSILLAOEKLKALDQCIYGACYSIEPLDLPQIQR 2885  
Qy 1861 HGLSAPSLHSYSPGSEINRVASCLKRGVPLRVVHRHARSVRARLLSQGGRAAATCGKYL 1920  
Db 2886 HGLSAPSLHSYSPGSEINRVASCLKRGVPLRVVHRHARSVRARLLSQGGRAAATCGKYL 2945  
Qy 1921 NNAVRTKLTPIPAASQLDLSFVAGYSGDDIYHSLSRARPFRWFMCLLLSVGVGIY 1980  
Db 2946 NNAVRTKLTPIPAASQLDLSFVAGYSGDDIYHSLSRARPFRWFMCLLLSVGVGIY 3005  
Qy 1981 LLPNR 1985  
Db 3006 LLPNR 3010  
RESULT 2  
Q9QP06 PRELIMINARY; PRT; 3010 AA.  
AC Q9QP06;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polypeptide.  
OS Hepatitis C virus type 1b.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 1.  
OX NCBI\_TaxID=31647;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=99370154; PubMed=10438800;  
RX Koch J.O., Bartenschlager R.;  
RA "Modulation of hepatitis C virus NS5A hyperphosphorylation by  
RT nonstructural proteins NS3, NS4A, and NS4B.";  
RL J. Virol. 73:7138-7146(1999).  
DR EMBL: AJ238800; CAB53095.1; -  
DR PIR: A61196; A61196.  
DR PIR: PQ0246; PQ0246.  
DR PIR: FS0329; FS0329.  
DR HSP; P27958; IALR.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0019028; C: viral capsid; IEA.  
DR GO: GO:0019031; C: viral envelope; IEA.  
DR GO: GO:0005524; F: ATP binding; IEA.  
DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.  
DR GO: GO:0003723; F: RNA binding; IEA.  
DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.  
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.  
DR GO: GO:0006350; P: transcription; IEA.  
DR GO: GO:0019079; P: viral genome replication; IEA.  
DR GO: GO:0019087; P: viral translation; IEA.  
DR InterPro: IPR000345; CytC\_heme\_BS.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR002522; HCV capsid.  
DR InterPro: IPR002521; HCV core.  
DR InterPro: IPR002519; HCV env.  
DR InterPro: IPR002531; HCV NS1.  
DR InterPro: IPR000745; HCV NS4A.  
DR InterPro: IPR001490; HCV NS4B.  
DR InterPro: IPR002868; HCV NS5A.  
DR InterPro: IPR002166; HCV RdRp.  
DR InterPro: IPR001650; Helicase C.  
DR InterPro: IPR004109; Peptidase\_S29.

DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro: IPR007095; RNA pol\_Ds\_PS.  
DR InterPro: IPR007094; RNA pol\_PSVir.  
DR Pfam: PF01543; HCV\_capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NS1; 1.  
DR Pfam: PF01538; HCV\_NS2; 1.  
DR Pfam: PF02907; HCV\_NS3; 1.  
DR Pfam: PF01006; HCV\_NS4a; 1.  
DR Pfam: PF01003; HCV\_NS4b; 1.  
DR Pfam: PF01506; HCV\_NS5a; 1.  
DR Pfam: PF00271; Helicase\_C; 1.  
DR Pfam: PF00998; Viral\_RdRp; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT CHAIN 810 1026 NS3 proteinase.  
FT CHAIN 1027 1657 NS3/4A proteinase cofactor.  
FT CHAIN 1658 1712 core protein.  
FT CHAIN 1 191 NS4B protein.  
FT CHAIN 1712 1972 NS5A phosphoprotein.  
FT CHAIN 1973 2419 NS5B RNA dependant RNA polymerase.  
FT CHAIN 2419 3010 glycoprotein E1.  
FT CHAIN 192 383 glycoprotein E2.  
FT CHAIN 384 746 p7 Peptide.  
FT CHAIN 747 809  
SQ SEQUENCE 3010 AA; 327000 MW; A570B9890DD64634 CRC64;  
Query Match 99.0%; Score 10365; DB 2; Length 3010;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1964; Conservative 8; Mismatches 13; Indels 0; Gaps 0;  
Qy 1 MAPITAYSQQTGLGLGCIITSLTGRDRNOVEGEVQVSTATQSLATCVNGVCVTVHGA 60  
Db 1026 LAHITAYSQQTGLGLGCIITSLTGRDRNOVEGEVQVSTATQSLATCVNGVCVTVHGA 1085  
Qy 61 GSKTLAGPKPIITQMTYNVDQDLVGWQAPPQAKSLTPTCTGSSDLYLVTRHADVIPVRR 120  
Db 1086 GSKTLAGPKPIITQMTYNVDQDLVGWQAPPQAKSLTPTCTGSSDLYLVTRHADVIPVRR 1145  
Qy 121 GDSRGSLLSPRVSYLKGSSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 1146 GDSRGSLLSPRVSYLKGSSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
Qy 181 RSPVFTDSSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLSPG 240  
Db 1206 RSPVFTDSSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLSPG 1265  
Qy 241 AYMSKAHGIDPNIRTVRTITTCAPTITSTYTGKFLADGGCGGAYDIIICDECHSTDTT 300  
Db 1266 AYMSKAHGIDPNIRTVRTITTCAPTITSTYTGKFLADGGCGGAYDIIICDECHSTDTT 1325  
Qy 301 ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNPNEEVALSSTGEPYGAIPETI 360  
Db 1326 ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNPNEEVALSSTGEPYGAIPETI 1385  
Qy 361 KGRHLIFCHSKKKCKDELAALKSLGLNAVAYRGLDVSVPTSGDVIVVATDALTGTFT 420  
Db 1386 KGRHLIFCHSKKKCKDELAALKSLGLNAVAYRGLDVSVPTSGDVIVVATDALTGTFT 1445  
Qy 421 GPDFSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRQRGRGRGMGIYRVTPG 480  
Db 1446 GPDFSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRQRGRGRGMGIYRVTPG 1505  
Qy 481 ERPSGMFDSVLCCEYDAGCWAYELTPAETSRLRAYLNTPLGIPVQDHLFEWESVFTGL 540  
Db 1506 ERPSGMFDSVLCCEYDAGCWAYELTPAETSRLRAYLNTPLGIPVQDHLFEWESVFTGL 1565  
Qy 541 THIDAHFLSQTQAGDNFFYLVAQYATVCARAAQAPPPSDQWKKLIRLKLPTLHGFTPL 600

1566 THIDAHFISQTKAQGNRPYLVAQATVCARAQAPPPSWDQWKKLRLKPTLHGPTPL 1625  
 601 YRLGAVQNEVTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 660  
 1626 YRLGAVQNEVTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 1685  
 661 IILSKPAIIPREVLYREFDEMEECASHLPYIEQGMQALAEQFKQKAIIGLQATKQAEA 720  
 1686 IIVLSKPAIIPREVLYREFDEMEECASHLPYIEQGMQALAEQFKQKAIIGLQATKQAEA 1745  
 721 AAPVVEKWRTEAFWAKEMWFIISGIVLAGLSTLPGNPAIASLMAFTASITSPITTOH 780  
 1746 AAPVVEKWRTEAFWAKEMWFIISGIVLAGLSTLPGNPAIASLMAFTASITSPITTOH 1805  
 781 TLLFNILGWAQAALAPPAAAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 840  
 1806 TLLFNILGWAQAALAPPAAAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 1865  
 841 FKVMSEMESTEDLVNLLPAILSPGALVGVVCAATLRRHVGPGEAGVQWMLIAFASR 900  
 1866 FKVMSEMESTEDLVNLLPAILSPGALVGVVCAATLRRHVGPGEAGVQWMLIAFASR 1925  
 901 GNVSPTHVPESSDAARVTOILSSLTITQLLKRLHOWINEDCTPCSGSWLRDVMWIC 960  
 1926 GNVSPTHVPESSDAARVTOILSSLTITQLLKRLHOWINEDCTPCSGSWLRDVMWIC 1985  
 961 TVLTFKTLQSKLLPRLPGVPPFCQYKGVWGRGIMQTTCPGCAQITGHVKNQSMR 1020  
 1986 TVLTFKTLQSKLLPRLPGVPPFCQYKGVWGRGIMQTTCPGCAQITGHVKNQSMR 2045  
 1021 IVGPRTCSNTHGTFPINAYTTCPTSPAPNYSRALNRVAEYVEVTRVGDVHVTGM 1080  
 2046 IVGPRTCSNTHGTFPINAYTTCPTSPAPNYSRALNRVAEYVEVTRVGDVHVTGM 2105  
 1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE 1140  
 2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE 2165  
 1141 PDVAVLTSMITDPSHTAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHDSDPAD 1200  
 2166 PDVAVLTSMITDPSHTAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHDSDPAD 2225  
 1201 LIEANLLWQEMGNITRVSEENKVLDSFEPLQAEEDREVSVPAAEILRRSRKPPRAM 1260  
 2226 LIEANLLWQEMGNITRVSEENKVLDSFEPLQAEEDREVSVPAAEILRRSRKPPRAM 2285  
 1261 PIWAPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSALA 1320  
 2286 PIWAPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSALA 2345  
 1321 ELATKTFGSSSAVDSGTATAPDQPSDDGAGSDVESYSSMPPLEGEFGDPLSDGSW 1380  
 2346 ELATKTFGSSSAVDSGTATAPDQPSDDGAGSDVESYSSMPPLEGEFGDPLSDGSW 2405  
 1381 STVSEERASDVCCSMSTYTGALITPCAAEETKLPINALSNLRRHNLVYATTSAS 1440  
 2406 STVSEERASDVCCSMSTYTGALITPCAAEETKLPINALSNLRRHNLVYATTSAS 2465  
 1441 LRQKKVTFDLQVLDHYRDVLEKMKAKASTVKAKLSVEEACKLTPPHSARSKFGYCAK 1500  
 2466 LRQKKVTFDLQVLDHYRDVLEKMKAKASTVKAKLSVEEACKLTPPHSARSKFGYCAK 2525  
 1501 DVNLSKAVNHIRSVWKKLLEDETETPIDTITMAKNEVFCVQPEKGRKPARLIVFDL 1560  
 2526 DVNLSKAVNHIRSVWKKLLEDETETPIDTITMAKNEVFCVQPEKGRKPARLIVFDL 2585  
 1561 VRUCEKMAVYVSTLPOAVMGSSVGFQYSPQGVFLVNAWAKKCPMGFAYDTRCFDS 1620  
 2586 VRUCEKMAVYVSTLPOAVMGSSVGFQYSPQGVFLVNAWAKKCPMGFAYDTRCFDS 2645  
 1621 TVTENDIRVEESIYQCCDLAEPARQAIRSLTERLYIGGLPNSKQNGCYRRRCASGVL 1680  
 2646 TVTENDIRVEESIYQCCDLAEPARQAIRSLTERLYIGGLPNSKQNGCYRRRCASGVL 2705

1681 TSCGNLTCTYLKAAACRAAKLQDCTMLVCGDDLVVICSAGTQDEASIRAFTEAMTRY 1740  
 2706 TSCGNLTCTYLKAAACRAAKLQDCTMLVCGDDLVVICSAGTQDEASIRAFTEAMTRY 2765  
 1741 SAPPDGPPEVDLELITSCSSNVSAHDASGRVYVLTTRDPTTPLARAWEETARHTPVN 1800  
 2766 SAPPDGPPEVDLELITSCSSNVSAHDASGRVYVLTTRDPTTPLARAWEETARHTPVN 2825  
 1801 SWLGNIMYAPTLMWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDQIQL 1860  
 2826 SWLGNIMYAPTLMWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDQIQL 2885  
 1861 HGLSAFSLHSYSGEINRVASCLRLKLGVEPLRVWRHRSVRARLLSQGGAATCGKLYF 1920  
 2886 HGLSAFSLHSYSGEINRVASCLRLKLGVEPLRVWRHRSVRARLLSQGGAATCGKLYF 2945  
 1921 NNAVRTKLTPIPAASQDLSSWVAGYSGGDIYHLSRARPWFMCWLLLSVGVGIY 1980  
 2946 NNAVRTKLTPIPAASQDLSSWVAGYSGGDIYHLSRARPWFMCWLLLSVGVGIY 3005  
 1981 LLPNR 1985  
 3006 LLPNR 3010

RESULT 3

Q9J3H2 PRELIMINARY; PRT; 3010 AA.  
 AC Q9J3H2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 CC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MD20;  
 RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 EMBL: AF207761; AAF65951.1;  
 PIR: A61196; A61196.  
 PIR: PQ0246; PQ0246.  
 PIR: PS0329; PS0329.  
 DR HSP; P26663; IQUV.  
 DR GO: 0016021; C: integral to membrane; IEA.  
 DR GO: 0019028; C: viral capsid; IEA.  
 DR GO: 0019031; C: viral envelope; IEA.  
 DR GO: 0005824; F: ATP binding; IEA.  
 DR GO: 0008026; F: ATP-dependent helicase activity; IEA.  
 DR GO: 0003723; F: RNA binding; IEA.  
 DR GO: 0003968; F: RNA-directed RNA polymerase activity; IEA.  
 DR GO: 0008236; F: serine-type peptidase activity; IEA.  
 DR GO: 0005198; F: structural molecule activity; IEA.  
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.  
 DR GO: 0006350; P: transcription; IEA.  
 DR GO: 0019079; P: viral genome replication; IEA.  
 DR GO: 0019087; P: viral transformation; IEA.  
 DR InterPro: IPR00345; CytC heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR02522; HCV capsid.  
 DR InterPro: IPR02521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRp.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR004109; Peptidase\_S29.

DR InterPro; IP0009003; Pept. Ser. Cys.  
DR InterPro; IP0002518; Pept\_U39\_HCV\_Ns2.  
DR InterPro; IP0007095; RNA\_pol\_DS\_PS.  
DR InterPro; IP0007094; RNA\_pol\_PSVir.  
DR Pfam; PF01543; HCV\_capsid; 1.  
DR Pfam; PF01542; HCV\_core; 1.  
DR Pfam; PF01539; HCV\_env; 1.  
DR Pfam; PF01560; HCV\_NS1; 1.  
DR Pfam; PF01538; HCV\_NS2; 1.  
DR Pfam; PF02907; HCV\_NS3; 1.  
DR Pfam; PF01006; HCV\_NS4a; 1.  
DR Pfam; PF01001; HCV\_NS4b; 1.  
DR Pfam; PF01506; HCV\_NS5a; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR Pfam; PF00998; Viral\_RdRP; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
SQ SEQUENCE 3010 AA; 326764 MW; 1A48EB4B5E1440D0 CRC64;  
  
Query Match 97.9%; Score 10245; DB 2; Length 3010;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1929; Conservative 30; Mismatches 26; Indels 0; Gaps 0;  
  
QY 1 MAPITAYSQOTRGLGCIITSLTGRDNOVEGEVQVYSTATQSLATCVNGVCWTVYHGA 60  
DB 1026 LAPITAYAQOTRGLGCIITSLTGRDKNOVEGEVQVYSTATQSLATCVNGVCWTVYHGA 1085  
  
QY 61 GSKTLGAPKGIITQMYTNVDQVLGVQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 120  
DB 1086 GSKALAGQKPVQMYTNVDQVLGVQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 1145  
  
QY 121 GDSRGLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
DB 1146 GDSRGLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
  
QY 181 RSPVFTDSSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGVKVLVNPVAAITLFGF 240  
DB 1206 RSPVFTDSSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGVKVLVNPVAAITLFGF 1265  
  
QY 241 AYMSKAGHDIPNIRTVGRTITTCAPITYSYGKFLADGGCGGGAYDIIICDECHSTDSTT 300  
DB 1266 AYMSKAGHDVPNIRTVGRTITTCAPITYSYGKFLADGGCGGGAYDIIICDECHSTDSTT 1325  
  
QY 301 ILGIGTVLDOAETAGARLVVLAATPPGSVTVPHNPTEEVVALSGTGEIPYKAIPIETI 360  
DB 1326 ILGIGTVLDOAETAGARLVVLAATPPGSVTVPHNPTEEVVALSGTGEIPYKAIPIETI 1385  
  
QY 361 KGRGHILFCHSKKKCKDELAALKSLGILNAVAYYRGLDVSVIPTSGDVIIVVATDALTMTGFT 420  
DB 1386 KGRGHILFCHSKKKCKDELAALKSLGILNAVAYYRGLDVSVIPTSGDVIIVVATDALTMTGFT 1445  
  
QY 421 GPDFSDVDCNTCVTQVDFSLDPTFTIETTTPQDAVRSRQRRGTRGRGMGIYRFVTPG 480  
DB 1446 GPDFSDVDCNTCVTQVDFSLDPTFTIETTTPQDAVRSRQRRGTRGRGMGIYRFVTPG 1505  
  
QY 481 ERPSGMFDSVLCEDYDAGCANYELTPAETSRLRAYLNTPLPVCOHDLHLEFWESVFTGL 540  
DB 1506 ERPSGMFDSVLCEDYDAGCANYELTPAETSRLRAYLNTPLPVCOHDLHLEFWESVFTGL 1565  
  
QY 541 THIDAHFLSOTKAGDNFPVLVAYQATVCARAQAPPSWDMWKCLIRLXPTLHGTPPLL 600  
DB 1566 THIDAHFLSOTKAGDNFPVLVAYQATVCARAQAPPSWDMWKCLIRLXPTLHGTPPLL 1625  
  
QY 601 YRLGAVQNEVTHPTIKYIMACMSADLEVVVSTWLVGGVLAALAAAYCLITTSGVVIVGR 660  
DB 1626 YRLGAVQNEVTLTHPTIKYIMACMSADLEVVVSTWLVGGVLAALAAAYCLITTSGVVIVGR 1685  
  
QY 661 IILSGKPAIIPDREVLRYEFDEMEECASHLPYIEQGMQLAEQFKQKAGILLOTATKQAEA 720  
DB 1686 IILSGKPAVIPDRDVLRYEFDEMEECASHLPYIEQGMQLAEQFKQKAGILLOTATKQAEA 1745

QY 721 AAPVVESKWTLEAFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSLPTTQH 780  
DB 1746 AAPVVESKWALETFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSLPTTQH 1805  
  
QY 781 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
DB 1806 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1865  
  
QY 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAVQMMNRLIAFASR 900  
DB 1866 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAVQMMNRLIAFASR 1925  
  
QY 901 GNVSPHYYPESDAAARVTQILSSLTITQLKELHOWINEDCSTPCSGSLWRDWDWIC 960  
DB 1926 GNVSPHYYPESDAAARVTQILSSLTITQLKELHOWINEDCSTPCSGSLWRDWDWIC 1985  
  
QY 961 TVLTDFTKWLQSKLLPRLPGVPFFSCORGKGVKWRGDMGIMOTTCPCGAOITGHVKNQSMR 1020  
DB 1986 AVLTDFTKWLQSKLLPRLPGVPFFSCORGKGVKWRGDMGIMOTTCPCGAOITGHVKNQSMR 2045  
  
QY 1021 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 1080  
DB 2046 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 2105  
  
QY 1081 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPLLEBEVTFVLGNLYVLSOLPCEPE 1140  
DB 2106 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPLLEBEVTFVLGNLYVLSOLPCEPE 2165  
  
QY 1141 PDVAVLITSLMTDPSHITAEAKRLARGSPSPSLASSASOLSPSLKATCTTRHDSPAD 1200  
DB 2166 PDVAVLITSLMTDPSHITAEAKRLARGSPSPSLASSASOLSPSLKATCTTRHDSPAD 2225  
  
QY 1201 LIPANLLWRQEMGNI TRVESENKWLDSFEPLQAEDEEREVSVAEILRKRKFFRAM 1260  
DB 2226 LIPANLLWRQEMGNI TRVESENKWLDSFEPLQAEDEEREVSVAEILRKRKFFRAM 2285  
  
QY 1261 PIWARPDPNPPLLESWKDPDYVPVHGCPLPAKAPPDPPRRKRKTVLSESTVSALA 1320  
DB 2286 PIWARPDPNPPLLESWKDPDYVPVHGCPLPAKAPPDPPRRKRKTVLSESTVSALA 2345  
  
QY 1321 ELATKTFGSESAVDSGTATSDPDSDDGSDGSDVESYSSMPLEGEPPDLSGWS 1380  
DB 2346 ELATKTFGSESAVDSGTATSDPDSDDGSDGSDVESYSSMPLEGEPPDLSGWS 2405  
  
QY 1381 STVSEBASEDVVCCSMSTWTGALITPCAABETKLIPINALNSLLRHHNIVYATTSRSAS 1440  
DB 2406 STVSEBASEDVVCCSMSTWTGALITPCAABETKLIPINALNSLLRHHNIVYATTSRSAS 2465  
  
QY 1441 LROKKYTFDELQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
DB 2466 LROKKYTFDELQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
  
QY 1501 DVNLSKSAVNHIRSVMKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDDL 1560  
DB 2526 DVNLSKSAVNHIRSVMKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDDL 2585  
  
QY 1561 VRVCEKVALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAMKAKCPMGFAYDTRCFDS 1620  
DB 2586 VRVCEKVALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAMKAKCPMGFAYDTRCFDS 2645  
  
QY 1621 TVTENDIRVEESIYQCDDIAPEARQAIRSLTERLYTGGPLTNSKGNCGYRRRCRASGVLT 1680  
DB 2646 TVTENDIRVEESIYQCDDIAPEARQAIRSLTERLYTGGPLTNSKGNCGYRRRCRASGVLT 2705  
  
QY 1681 TSCGNTLTCLYKAAAACRAAKLQDCTMLVCGDDLVIIVICESAGTOEDEASIRAFETAMTRY 1740  
DB 2706 TSCGNTLTCLYKAAAACRAAKLQDCTMLVCGDDLVIIVICESAGTOEDEASIRAFETAMTRY 2765  
  
QY 1741 SAPPDGPPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPIARAAMETARITPVN 1800  
DB 2766 SAPPDGPPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPIARAAMETARITPVN 2825  
  
QY 1801 SWLGNLIMYAPTILWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIQRL 1860

```
Db 2826 SWLGNINMYAPLWARMIMTHFFSILLAQEQLKALDQCIYACYSIEPLDLPQIQL 2885
QY 1861 HGLSAFSLHSYSPGBINRVASCLRLKGLVPLRWRHRARSVRARLLSQGGRATCGKLYF 1920
Db 2886 HGLSAFSLHSYSPGBINRVASCLRLKGLVPLRWRHRARSVRARLLSQGGRATCGKLYF 2945
QY 1921 NNAVETKLTIPPAASQLDLSWVAGYSGGDIYHSISRARPWFMCILLLSVGVGY 1980
Db 2946 NNAVETKLTIPPAASHLDLSWVAGYSGGDIYHSISRARPWFMCILLLSVGVGY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 4
Q9J3F4
ID Q9J3F4 PRELIMINARY; PRT; 3008 AA.
AC Q9J3F4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
GN Name=MD34;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD34;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208024; AAF61205.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; P26663; IJXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS4.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
```

```
DR PF02907; HCV NS3; 1.
DR PF01006; HCV NS4a; 1.
DR PF01001; HCV NS4b; 1.
DR PF01506; HCV NS5a; 1.
DR PF00271; Helicase C; 1.
DR PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3008 AA; 326835 MW; 99AE09E14C3109F4 CRC64;

Query Match 97.7%; Score 10225; DB 2; Length 3008;
Best Local Similarity 97.1%; Pred No. 0;
Matches 1928; Conservative 33; Mismatches 24; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFILATCNGVCWTYYHGA 60
Db 1024 LAPITAYSQOTRGLGCIITSLTGRDRKNQVEGEVQVSTATQSFILATCNGVCWTYYHGA 1083

QY 61 GSKTLAGPKGPIOMYTNVDQDLVQWQAPPGARSITPCTCGSSDLYLVTRHADVIPVRR 120
Db 1084 GSKTLAGPKGPIOMYTNVDQDLVQWQAPPGARSITPCTCGSSDLYLVTRHADVIPVRR 1143

QY 121 GDSRGLSPRPVSVYLGSSGGPLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db 1144 GDSRGLSPRPVSVYLGSSGGPLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 1203

QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKTKVPAAYAAQYKVLVLPNSVAATLFG 240
Db 1204 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKTKVPAAYAAQYKVLVLPNSVAATLFG 1263

QY 241 AYMSKANGIDNRITGVRTITTTGAPITYSTYCKFLADGCGSGAGVDIICDCSHSDSTT 300
Db 1264 AYMSKANGIDNRITGVRTITTTGAPITYSTYCKFLADGCGSGAGVDIICDCSHSDSTT 1323

QY 301 ILGIGTVLDQAEATAGARLVLATATATPPGVSIVPHENIEVALSSTGEIPFYKAIPETI 360
Db 1324 ILGIGTVLDQAEATAGARLVLATATATPPGVSIVPHENIEVALSSTGEIPFYKAIPETI 1383

QY 361 KGRHLIFCHSKKKCDLAALKSLGLNNAVAYRGLDVSVIPTSGDIVVATDALTMTGT 420
Db 1384 KGRHLIFCHSKKKCDLAALKSLGLNNAVAYRGLDVSVIPTSGDIVVATDALTMTGT 1443

QY 421 GDFDSVIDNCVCTQVDFSLDPTTITFTTTPQDAVSRSORRGTRGCRMGIRYRVTPG 480
Db 1444 GDFDSVIDNCVCTQVDFSLDPTTITFTTTPQDAVSRSORRGTRGCRMGIRYRVTPG 1503

QY 481 ERPSGMFSSVLCVCDAGCANVELTPAETS VRLRAYLNTPLGVPQDHLFEWESVFTGL 540
Db 1504 ERPSGMFSSVLCVCDAGCANVELTPAETS VRLRAYLNTPLGVPQDHLFEWESVFTGL 1563

QY 541 THIDAHFTLSQTKAGDNFPYLVAQATVCARQAQPPSDWQWKKLIRLKPITHGTPLL 600
Db 1564 THIDAHFTLSQTKAGDNFPYLVAQATVCARQAQPPSDWQWKKLIRLKPITHGTPLL 1623

QY 601 YRLGAVQNEVTTHTPIITKYIMACMSADLEVVTSWLVGVGLAALAAAYCLTTSVVIVGR 660
Db 1624 YRLGAVQNEVTTHTPIITKYIMACMSADLEVVTSWLVGVGLAALAAAYCLTTSVVIVGR 1683

QY 661 IILSGKPAIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKOKAIGLLQATKQABA 720
Db 1684 IILSGKPAIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKOKAIGLLQATKQAB 1743

QY 721 AAPVVEKWRITLEAFWAKHMWNFISGIQYLAGISTILPGNPAISLMAFTASITSLTTOH 780
Db 1744 AAPVVEKWRITLEAFWAKHMWNFISGIQYLAGISTILPGNPAISLMAFTASITSLTTOH 1803

QY 781 TLLFNILGWAAQLAPPSSAASAFVAGIAGAVSGISGLKVLVDILAGYAGAGALVA 840
Db 1804 TLLFNILGWAAQLAPPSSAASAFVAGIAGAVSGISGLKVLVDILAGYAGAGALVA 1863

QY 841 FKVMGEMPSTEDLVNLLPAILSPGALVVGVCVCAAILRRHVGEVGMNRLIAPSR 900
```

1864 FKVMSGEMSTEDLVNLLPAILSPGALVGVCAAILRRVVGOGGAVQNMNLLIAPSR 1923  
901 GNVHSTHYVPESDAAARVTOILSSLTITQLLKLHQMINEDCSTPCSGSWLRDWDWIC 960  
1924 GNVHSTHYVPESDAAARVTOILSSLTITQLLKLHQMINEDCSTPCSGSWLRDWDWIC 1983  
961 TVLTDFKTLWLOSLRLPLPGVPFSCORGKYGWVRGDMQTTCPGCAQITGHVKGSMR 1020  
1984 TVLTDFKTLWLOSLRLPLPGVPFSCORGKYGWVRGDMQTTCPGCAQITGHVKGSMR 2043  
1021 IUGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALRWVAABEYEVTVRGDFHYVTGM 1080  
2044 IUGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALRWVAABEYEVTVRGDFHYVTGM 2103  
1081 TTDNVKCPQVPAPEFTEVDGVRHLHRYAPACKPLLRREEVTVLGLNQYLVGSQLPCEPE 1140  
2104 TTDNVKCPQVPAPEFTEVDGVRHLHRYAPACKPLLRREEVTVLGLNQYLVGSQLPCEPE 2163  
1141 POVAVLTSMITDPSHITAEAKRLARGSPSPSLASSASQLSAPSLKATCTTTHDGFDDAD 1200  
2164 POVAVLTSMITDPSHITAEAKRLARGSPSPSLASSASQLSAPSLKATCTTTHDGFDDAD 2223  
1201 LIEANLLWRQEMGNGNITRVESENKVVILDSFELQAEDEDEREVSVPAAELRRRKRKPPRAM 1260  
2224 LIEANLLWRQEMGNGNITRVESENKVVILDSFELQAEDEDEREVSVPAAELRRRKRKPPRAM 2283  
1261 PIWARDYNPPLLESKWDPDYPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
2284 PIWARDYNPPLLESKWDPDYPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 2343  
1321 ELATKTFGSSSAVDSGTATASPDQSDSDGSDVESYSSMPPLLEGEGDPLDSDGSW 1380  
2344 ELATKTFGSSSAVDSGTATASPDQSDSDGSDVESYSSMPPLLEGEGDPLDSDGSW 2403  
1381 STVSEASDSDVCCSMYSYTWGALITPCAAETKLPINALNSLLRHNLVYATTSRSAS 1440  
2404 STVSEASDSDVCCSMYSYTWGALITPCAAETKLPINALNSLLRHNLVYATTSRSAS 2463  
1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
2464 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 2523  
1501 DVNLSKAVNHRSVWKOLLETTETPIDTIMAKNEVFCVQKEGKGRKPARLIVFPDLG 1560  
2524 DVNLSKAVNHRSVWKOLLETTETPIDTIMAKNEVFCVQKEGKGRKPARLIVFPDLG 2583  
1561 VRVCERWALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
2584 VRVCERWALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 2643  
1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYTGGPVTNSKQNGCYRRCRASGVLT 1680  
2644 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYTGGPVTNSKQNGCYRRCRASGVLT 2703  
1681 TSCGNLTCTYLKAAACRAKLODCTMLVCGDDLVVICESAGTQDEASLRATEAMTRY 1740  
2704 TSCGNLTCTYLKAAACRAKLODCTMLVCGDDLVVICESAGTQDEASLRATEAMTRY 2763  
1741 SAPPDGPPEYDLELITSCSNVSVAHADSGKRVYVLTDRPTTPLARAAWETARTTPVN 1800  
2764 SAPPDGPPEYDLELITSCSNVSVAHADSGKRVYVLTDRPTTPLARAAWETARTTPVN 2823  
1801 SWLGNITMVAPTLWARMILMTHFFSILLAQEALDQCIYGCYSIEPLDLPQIIQRL 1860  
2824 SWLGNITMVAPTLWARMILMTHFFSILLAQEALDQCIYGCYSIEPLDLPQIIQRL 2883  
1861 HGLSASLSHSYSGEINRVASCLRKLGVPPLRVHRARSVRARLLSQGGRAATCGKLYF 1920  
2884 HGLSASLSHSYSGEINRVASCLRKLGVPPLRVHRARSVRARLLSQGGRAATCGKLYF 2943  
1921 NNAVRTKLTPTIPAAASQLDSNFWAGYGGDIYHLSRARPWFWMCLLLLSVGVGIY 1980

Db 2944 NNAVRTKLTPTIPAAASQLDSNFWAGYGGDIYHLSRARPWFWMCLLLLSVGVGIY 3003  
QY 1981 LLENR 1985  
Db 3004 LLENR 3008  
RESULT 5  
QBV638 PRELIMINARY; PRT; 3010 AA.  
AC QBV638;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fanning L.J., Itakura J., Nagayama K., Enomoto N.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF313916; AAL55821.1;  
DR PIR: A61196; A61196.  
DR PIR: PQ0246; PQ0246.  
DR PIR: PQ0804; PQ0804.  
DR PIR: PS0329; PS0329.  
DR GO: GO:0016021; C:integral to membrane; IEA.  
DR GO: GO:0019028; C:viral capsid; IEA.  
DR GO: GO:0019031; C:viral envelope; IEA.  
DR GO: GO:0005524; F:ATP binding; IEA.  
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO: GO:0003723; F:RNA binding; IEA.  
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO: GO:0005198; F:structural molecule activity; IEA.  
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.  
DR GO: GO:0006350; P:transcription; IEA.  
DR GO: GO:0019079; P:viral genome replication; IEA.  
DR GO: GO:0019087; P:viral transformation; IEA.  
DR InterPro: IPR000345; CytC\_heme\_BS.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002521; HCV\_core.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NS1.  
DR InterPro: IPR000745; HCV\_NS4a.  
DR InterPro: IPR001490; HCV\_NS4b.  
DR InterPro: IPR002868; HCV\_NS5a.  
DR InterPro: IPR002166; HCV\_RdRP.  
DR InterPro: IPR001650; Helicase\_C.  
DR InterPro: IPR004109; Peptidase\_S29.  
DR InterPro: IPR009003; Pept Ser Cys.  
DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR007094; RNA\_pol\_Psvir.  
DR Pfam: PF01543; HCV\_capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NS1; 1.  
DR Pfam: PF01538; HCV\_NS2; 1.  
DR Pfam: PF02907; HCV\_NS3; 1.  
DR Pfam: PF01006; HCV\_NS4a; 1.  
DR Pfam: PF01001; HCV\_NS4b; 1.  
DR Pfam: PF01506; HCV\_NS5a; 1.  
DR Pfam: PF00271; Helicase\_C; 1.  
DR Pfam: PF00998; Viral\_RdRP; 1.  
DR SMART: SM00487; DEXDG; 1.  
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 3010 3010



SQ	SEQUENCE	3010 AA; 327182 MW; 33AAA6C07251C839 CRC64;	
	Query Match	97.7%; Score 10225; DB 2; Length 3010;	
	Best Local Similarity	97.2%; Pred. No. 0;	
	Matches 1930; Conservative	29; Mismatches 26; Indels 0; Gaps 0;	
QY	1	MAPITAYSOOTRGLGCIITSUTGRDRNOVEGEVWVSTATOSFLATCVNGVCWTVYHGA 60	
DB	1026	LAPITAYAOOTRGLGCIITSUTGRDRNOVEGEVWVSTATOSFLATCVNGVCWTVYHGA 1085	
QY	61	GSKTLGPKGPIITQMTYNTVDQDLVGWQAPPGARSLTPTCTGSSDLYLVTRHADVIVRRR 120	
DB	1086	GKTLGAGKGPITQMTYNTVDQDLVGWLAAPPGARSLTPTCTGSSDLYLVTRHADVIVRRR 1145	
QY	121	GDSRGLSPRPVSVLKSSGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVSVESMETTM 180	
DB	1146	GDSRGLSPRPVSVLKSSGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVSVESMETTM 1205	
QY	181	RSPVFTDNSSPPAVPQTFOVAHLHAPTGSCKTKVPAAYAAQYKVLVLPNSVAATLGF 240	
DB	1206	RSPVFTDNSSPPAVPQTFOVAHLHAPTGSCKTKVPAAYAAQYKVLVLPNSVAATLGF 1265	
QY	241	AYMSKAHGDIPNIRTVGVRITITGAPITYSTYKFLADGCGSGAYDIIICDECHSTDST 300	
DB	1266	AYMSKAHGDIPNIRTVGVRITITGAPITYSTYKFLADGCGSGAYDIIICDECHSTDST 1325	
QY	301	ILGIGTVLDOASTAGARLVLTATPPGSVTVPHNIEVALSSTGEIPFYKKAIPETI 360	
DB	1326	ILGIGTVLDOASTAGARLVLTATPPGSVTVPHNIEVALSSTGEIPFYKKAIPETI 1385	
QY	361	KGGRHLIFCHSKKKCKDELAALKSGLGNVAVYRGLDVSVIPTSGDVIIVATDALMTGT 420	
DB	1386	KGGRHLIFCHSKKKCKDELAALKSGLGNVAVYRGLDVSVIPTSGDVIIVATDALMTGT 1445	
QY	421	GDPSVIDCNTCVTQTVDFSLDPTFTIETTVPQDAVSRSQRGRMGIIYFVPTG 480	
DB	1446	GDPSVIDCNTCVTQTVDFSLDPTFTIETTVPQDAVSRSQRGRMGIIYFVPTG 1505	
QY	481	EPSCGMFDSVLCVCEYDAGCANVELTPAETSRLRAYLNTRGIPVCOHLEWESVFTGL 540	
DB	1506	EPSCGMFDSVLCVCEYDAGCANVELTPAETSRLRAYLNTRGIPVCOHLEWESVFTGL 1565	
QY	541	THIDAHFLSQTQAGDNFYLVAQATVCARAQAAPPSPDWOMWKCLIRLKPILHGPPL 600	
DB	1566	THIDAHFLSQTQAGDNFYLVAQATVCARAQAAPPSPDWOMWKCLIRLKPILHGPPL 1625	
QY	601	YRLGAVQNEVTHPTIKTIFIMACMSADLEWVTSWVLVGGVLAALAAAYCLITGSSVIVGR 660	
DB	1626	YRLGAVQNEVTHPTIKTIFIMACMSADLEWVTSWVLVGGVLAALAAAYCLITGSSVIVGR 1685	
QY	661	IILSGKPALIPREVLYRFRFDEMEECASHLPVTEQMOQLAEOFKOKAIGLLOTATKQAEA 720	
DB	1686	IILSGKPALIPREVLYRFRFDEMEECASHLPVTEQMOQLAEOFKOKAIGLLOTATKQAEA 1745	
QY	721	AAPVVEKWRITLFAFWAKEMWNPISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 780	
DB	1746	AAPVVEKWRITLFAFWAKEMWNPISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 1805	
QY	781	TLLFNILGWAAQAALAPPAAAFVAGIAGAAVSGIGLKVLDILAGYAGAGALVA 840	
DB	1806	TLLFNILGWAAQAALAPPAAAFVAGIAGAAVSGIGLKVLDILAGYAGAGALVA 1865	
QY	841	EKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMKRLIAFASR 900	
DB	1866	EKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMKRLIAFASR 1925	
QY	901	GNHVSPTHVPESDAARVTQILSSLTITQLKRLHOWINEDCSTPCSGSWLRDWDWMC 960	
DB	1926	GNHVSPTHVPESDAARVTQILSSLTITQLKRLHOWINEDCSTPCSGSWLRDWDWMC 1985	
QY	961	TVLTFKTLQSKLLPRLPGVPFSCQYKGYWVRGDMQITPCGAQITGHVKNQSMR 1020	
DB	1986	TVLTFKTLQSKLLPRLPGVPFSCQYKGYWVRGDMQITPCGAQITGHVKNQSMR 2045	

QY	1021	IVGPETCSNTWHGTTPINAYTTGCTPSPAPNYSRALRVAAEEYVEVTRVGDHFVYTCM 1080	
DB	2046	IVGPETCSNTWHGTTPINAYTTGCTPSPAPNYSRALRVAAEEYVEVTRVGDHFVYTCM 2105	
QY	1081	TTDNYKCPQCPAPPEFFTEVDGVRHLRYAPACKPLREEVTEFLVGLNQYLVGSQPCBE 1140	
DB	2106	TTDNYKCPQCPAPPEFFTEVDGVRHLRYAPACKPLREEVTEFLVGLNQYLVGSQPCBE 2165	
QY	1141	PDVAVLTSMLTDPSSHITAEATAKRLARGSPSPSLASSASQLSAPSJKATCTTRHSDPAD 1200	
DB	2166	PDVAVLTSMLTDPSSHITAEATAKRLARGSPSPSLASSASQLSAPSJKATCTTRHSDPAD 2225	
QY	1201	LIEANLLWRQEMGNIITRVESKVVILDSPEPIQAEEDEREVSYPAEILRRSRKPPRAM 1260	
DB	2226	LIEANLLWRQEMGNIITRVESKVVILDSPEPIQAEEDEREVSYPAEILRRSRKPPRAM 2285	
QY	1261	PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPPIPPRRKRTVVLSSTVSSALA 1320	
DB	2286	PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPPIPPRRKRTVVLSSTVSSALA 2345	
QY	1321	ELATKTFGSSSSAVDSGTATASPOPSDDGDAGSDVESYSSMPLECEPGDPLSDGSM 1380	
DB	2346	ELATKTFGSSSSAVDSGTATASPOPSDDGDAGSDVESYSSMPLECEPGDPLSDGSM 2405	
QY	1381	STVSEEAAGEDVVCSSMSYTWGALITPCAABETKLPINALSNLLRHHNLVYATTSSAS 1440	
DB	2406	STVSEEAAGEDVVCSSMSYTWGALITPCAABETKLPINALSNLLRHHNLVYATTSSAS 2465	
QY	1441	LQKQVTFDRLOVLDHDDYDVLKEMKAKASTYKAKLLSVEEACKLTPPHSARSKFGYGA 1500	
DB	2466	LQKQVTFDRLOVLDHDDYDVLKEMKAKASTYKAKLLSVEEACKLTPPHSARSKFGYGA 2525	
QY	1501	DYRNISSKAVAHIRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGRKPAELIVFPDGL 1560	
DB	2526	DYRNISSKAVAHIRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGRKPAELIVFPDGL 2585	
QY	1561	VVCEKMAKYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFYDTRCFDS 1620	
DB	2586	VVCEKMAKYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFYDTRCFDS 2645	
QY	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRCRAGSVLT 1680	
DB	2646	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRCRAGSVLT 2705	
QY	1681	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDLVVICESAGTQDEASLRAFTAMTRY 1740	
DB	2706	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDLVVICESAGTQDEASLRAFTAMTRY 2765	
QY	1741	SAPPGDPPKPEYDLELITSCSSNVSAHDASGRVYVYLRDPTTTPPLARAAMETARHTPVN 1800	
DB	2766	SAPPGDPPKPEYDLELITSCSSNVSAHDASGRVYVYLRDPTTTPPLARAAMETARHTPVN 2825	
QY	1801	SWLGNIIYAPTLMWARMILMTHFFSILLAQEOLEKALDCQIYGCYCSIEPLDLPQILQRL 1860	
DB	2826	SWLGNIIYAPTLMWARMILMTHFFSILLAQEOLEKALDCQIYGCYCSIEPLDLPQILQRL 2885	
QY	1861	HGLSAPLSHSYSPGIRNVASCLKLYDPLRWHRARSVEARLLSOGGRAATCGKYL 1920	
DB	2886	HGLSAPLSHSYSPGIRNVASCLKLYDPLRWHRARSVEARLLSOGGRAATCGKYL 2945	
QY	1921	NNAVTKLKITPIPAASQDLSSWFVAGYSGGDIYHLSLRARPRWFMWMLLLLSVGVGY 1980	
DB	2946	NNAVTKLKITPIPAASQDLSSWFVAGYSGGDIYHLSLRARPRWFMWMLLLLSVGVGY 3005	
QY	1981	LLPNR 1985	
DB	3005	LLPNR 3010	

RESULT 6  
O93016 PRELIMINARY; PRT; 3010 AA.  
ID O93016

AC O93016;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98232263; PubMed=9572551;  
 RA Trowbridge R., Gowans B.J.;  
 RT "Molecular cloning of an Australian isolate of hepatitis C virus.";  
 RL Arch. Virol. 143:501-511(1998).  
 DR EMBL: AJ000009; CAA03854.1; -;  
 DR PIR: A61196; A61196.  
 DR PIR: P02446; P02446.  
 DR PIR: P02555; P02555.  
 DR PIR: P03299; P03299.  
 DR HSSP: P26663; JUXP.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; C: viral envelope; IEA.  
 DR GO: GO:0005524; F: ATP binding; IEA.  
 DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.  
 DR GO: GO:0003723; F: RNA binding; IEA.  
 DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.  
 DR GO: GO:0008236; F: serine-type peptidase activity; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.  
 DR GO: GO:0006350; P: transcription; IEA.  
 DR GO: GO:0019079; P: viral genome replication; IEA.  
 DR GO: GO:0019087; P: viral transformation; IEA.  
 DR InterPro: IPR000345; Cyt c heme\_Bs.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR002522; HCV capsid.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRP.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR004109; Peptidase\_S29.  
 DR InterPro: IPR009003; Pept Ser Cys.  
 DR InterPro: IPR002518; Pept U39 HCV NS2.  
 DR InterPro: IPR007095; RNA\_pol\_Ds\_P5.  
 DR InterPro: IPR007094; RNA\_pol\_P5vir.  
 DR Pfam: PF01543; HCV\_capsid; 1.  
 DR Pfam: PF01542; HCV\_core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 DR Pfam: PF01560; HCV NS1; 1.  
 DR Pfam: PF01538; HCV NS2; 1.  
 DR Pfam: PF02907; HCV NS3; 1.  
 DR Pfam: PF01006; HCV NS4a; 1.  
 DR Pfam: PF01001; HCV NS4b; 1.  
 DR Pfam: PF01506; HCV NS5a; 1.  
 DR Pfam: PF00271; Helicase\_C; 1.  
 DR Pfam: PF00998; Viral\_RdRP; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN\_1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 Polyprotein; Transmembrane.  
 FT CHAIN 810 1026 NS2 protein.  
 FT CHAIN 1027 1657 NS3 protein.  
 FT CHAIN 1658 1711 NS4a protein.  
 FT CHAIN 1 191 core protein.  
 FT CHAIN 1712 1972 NS4b protein.  
 FT CHAIN 1973 2419 NS5a protein.  
 FT CHAIN 2420 3010 NS5b protein.  
 FT CHAIN 192 383 E1 protein.

FT	CHAIN	384	746	E2 protein.
FT	CHAIN	747	809	p7 protein.
SQ	SEQUENCE	3010 AA;	327121 MW;	OEE02EDA54A8B61D CRC64;
	Query Match		97.6%;	Score 10218; DB 2; Length 3010;
	Best Local Similarity		97.0%;	Pred. No. 0;
	Matches 1925; Conservative		35; Mismatches	25; Indels 0; Gaps 0;

  

QY	1	MAPITAYSQQTGLGCGIITSLTGRDRNQVEGEVQVWVSTATQSFATFCVNGVCWTVVHGA	60
DB	1026	LAPITAYSQQTGLGCGIITSLTGRDKNQVEGEVQVWVSTATQSFATFCVNGVCWTVVHGA	1085
QY	61	GSKTLAGPKGPIOMYTNVDQDLVGHQAPPGARSILPCTCGSSDLXLVTRHADVIPVRRR	120
DB	1086	GSKTLAGPKGPVTOMYTNVDQDLVGHQAPPGARSILPCTCGSSDLXLVTRHADVIPVRRR	1145
QY	121	GDGRGSLSPRPVSYLKSGSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM	180
DB	1146	GDGRGSLSPRPVSYLKSGSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM	1205
QY	181	RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAATLFGF	240
DB	1206	RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAATLFGF	1265
QY	241	AYMSKAHGIDPNIRTVGVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	300
DB	1266	AYMSKAHGIDPNIRTVGVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	1325
QY	301	ILGIGTVLQDAETAGARLVVLTATPPGVSVPVPHPNIEEVALSSTGEIPIFYKAIPIETI	360
DB	1326	ILGIGTVLQDAETAGARLVVLTATPPGVSVPVPHPNIEEVALSSTGEIPIFYKAIPIETI	1385
QY	361	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGDVSVPITSGDVIWVATDALMTGFT	420
DB	1386	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGDVSVPITSGDVIWVATDALMTGFT	1445
QY	421	GDVDSVIDCNTCVTVQVDSLDPTFTIETTVQDAVSRSORGRTRGRMGYRFTVTPG	480
DB	1446	GDVDSVIDCNTCVTVQVDSLDPTFTIETTVQDAVSRSORGRTRGRMGYRFTVTPG	1505
QY	481	ERPDSMFDSVLCEDYDAGCANYELTPAETSVRRLAYLNTPLGVCDDHLEWESVFTGL	540
DB	1506	ERPDSMFDSVLCEDYDAGCANYELTPAETSVRRLAYLNTPLGVCDDHLEWESVFTGL	1565
QY	541	THIDAFHLSQTKOAGDNFFYLVAQATVCARQAQPPPSWDMQWKCLIRLKPTLHGPTPL	600
DB	1566	THIDAFHLSQTKOAGDNFFYLVAQATVCARQAQPPPSWDMQWKCLIRLKPTLHGPTPL	1625
QY	601	YRLGAVQNEVTHPTTKYIMACMSADLEWVSTWLVGGVLAALAAAYCLTTGSSVIVGR	660
DB	1626	YRLGAVQNEVTHPTTKYIMACMSADLEWVSTWLVGGVLAALAAAYCLTTGSSVIVGR	1685
QY	661	IILSGKPAIIPREVLYREFDEMEECASHLPYTEQGMQLAEQFKAKAIGLLQATKQAEA	720
DB	1686	IILSGKPAIIPREVLYREFDEMEECASHLPYTEQGMQLAEQFKAKAIGLLQATKQAEA	1745
QY	721	AAPVVEKWTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMFAFTASITSLTQH	780
DB	1746	AVPVVESKQWALEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMFAFTASITSLTQH	1805
QY	781	TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA	840
DB	1806	TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA	1865
QY	841	FKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAATLRRHVGPGEVAGVOMNRLTAFASR	900
DB	1866	FKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAATLRRHVGPGEVAGVOMNRLTAFASR	1925
QY	901	GNHVSPHYVPESDAAARVTQILSSITITQLKRLHQWINEDECSTPCSGSLRDVWDMTC	960
DB	1926	GNHVSPHYVPESDAAARVTQILSSITITQLKRLHQWINEDECSTPCSGSLRDVWDMTC	1985
QY	961	TVLITFKTLQSKLLPRLEPGVFFFSQQRQYKGVWRGDMQITQTCGCAQITGHVKNQSMR	1020

Db 1396 TVLDFKTLWLSKLLPRLPGIPFFYSQGYGKVRGDMQTTCGQAQIGHVANGSWR 2045  
QY 1021 IVGPRTCSNTWHGTGTPINAYTGTCTPSAPNYSRALWRVAAREYVEVTRVGDFFHYVTGM 1080  
Db 2046 IVGPRTCSNTWHGTGTPINAYTGTCTPSAPNYSRALWRVAAREYVEVTRVGDFFHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPPEFFTEVDGVRHRYAPACKPLREBTEVLGNGLYVCSQPCPE 1140  
Db 2106 TTDNVKCPQVPAPPEFFTEVDGVRHRYAPACKPLREBTEVLGNGLYVCSQPCPE 2165  
QY 1141 PDVAVLTSMLTDPDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTRHDSPAD 1200  
Db 2166 PDVAVLTSMLTDPDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWRQMGNNITRVESKNNVILDSFPLQAEEDERVSVPAILRRSRKFPFRAM 1260  
Db 2226 LIEANLLWRQMGNNITRVESKNNVILDSFPLQAEEDERVSVAAILRRSRKFPFRAM 2285  
QY 1261 PIWARPENPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPPRKRTVVLSSTVSSALA 1320  
Db 2286 PIWARPENPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPPRKRTVVLSSTVSSALA 2345  
QY 1321 ELATKTFGSSSESAVDSGTATASPDQSDGDAGSDVESYSSMPLEGEPPGDPDLSDGSW 1380  
Db 2346 ELATKTFGSSSESAVDSGTATASPDQSDGDAGSDVESYSSMPLEGEPPGDPDLSDGSW 2405  
QY 1381 STVSEASESDVVCSSMSYTWTCALITPCAAETKUPINALSNSLLRHNNLVATTSSRAS 1440  
Db 2406 STVSEASESDVVCSSMSYTWTCALITPCAAETKUPINALSNSLLRHNNLVATTSSRAS 2465  
QY 1441 LQKQKVTFRLOVLDHDDHVDLKEWAKASTVKALLSVEEACKLTPPHSARSKFGYGA 1500  
Db 2466 LQKQKVTFRLOVLDHDDHVDLKEWAKASTVKALLSVEEACKLTPPHSARSKFGYGA 2525  
QY 1501 DVNRLSSKAVNHIRSVWKKLLEDTETPIDTTIMAKNEFCVQPEKGRKPARLIIVFPDLG 1560  
Db 2526 DVNRLSSKAVNHIRSVWKKLLEDTETPIDTTIMAKNEFCVQPEKGRKPARLIIVFPDLG 2585  
QY 1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPQQRVEFLVNAWAKKCPMGFYDTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPQQRVEFLVNAWAKKCPMGFYDTRCFDS 2645  
QY 1621 TVTENDIRVERSIYOCDDLAPARQAIRSLTERLYTGGPLTNSKQNCQYRCRAGSVLT 1680  
Db 2646 TVTENDIRVERSIYOCDDLAPARQAIRSLTERLYTGGPLTNSKQNCQYRCRAGSVLT 2705  
QY 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICSAGTQEDASLRAFTAMTRY 1740  
Db 2706 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICSAGTQEDASLRAFTAMTRY 2765  
QY 1741 SAPPDGPPEYDLELITSCSNVSVVAHDASGRVYVLTTRDPTTLARAANETARTFPVN 1800  
Db 2766 SAPPDGPPEYDLELITSCSNVSVVAHDASGRVYVLTTRDPTTLARAANETARTFPVN 2825  
QY 1801 SWLGNIMVAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
Db 2826 SWLGNIMVAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIQRL 2885  
QY 1861 HGLSAFSLHSYSGEINRVASCLRLGVPLRVWRHARSVRLILSQGGAATCKGYLF 1920  
Db 2886 HGLSAFSLHSYSGEINRVASCLRLGVPLRVWRHARSVRLILSQGGAATCKGYLF 2945  
QY 1921 NNAVTKLKTPIPAASQDLDSWVAGYSGGDIYHLSRPRMFWCLILLSVGVGII 1980  
Db 2946 NNAVTKLKTPIPAASQDLDSWVAGYSGGDIYHLSRPRMFWCLILLSVGVGII 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

RESULT 7

Q9DTE7  
ID Q9DTE7 PRELIMINARY; PRT; 3010 AA.  
AC Q9DTE7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE=Serum;  
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,  
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,  
RA Mishiro S.;  
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
RT with hepatocellular carcinoma: the 'progression score' revisited.";  
RL Hepatol. Res. 20:161-171(2001).  
DR EMIL; AB049090; BAB18803.1; -;  
DR PIR; A61196; A61196.  
DR PIR; P00246; P00246.  
DR PIR; P00252; P00252.  
DR PIR; P00253; P00253.  
DR PIR; P00254; P00254.  
DR PIR; P00255; P00255.  
DR PIR; P00329; P00329.  
DR HSP; P26663; LUXP.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0019079; F:viral genome replication; IEA.  
DR GO; GO:0019087; F:viral transformation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002519; HCV\_env.  
DR InterPro; IPR002531; HCV\_NS1.  
DR InterPro; IPR000745; HCV\_NS4a.  
DR InterPro; IPR001490; HCV\_NS4b.  
DR InterPro; IPR002868; HCV\_NS5a.  
DR InterPro; IPR002166; HCV\_RdRP.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR004109; Peptidase\_S29.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_psvir.  
DR Pfam; PF01543; HCV\_capsid; 1.  
DR Pfam; PF01542; HCV\_core; 1.  
DR Pfam; PF01539; HCV\_env; 1.  
DR Pfam; PF01560; HCV\_NS1; 1.  
DR Pfam; PF01538; HCV\_NS2; 1.  
DR Pfam; PF02907; HCV\_NS3; 1.  
DR Pfam; PF01006; HCV\_NS4a; 1.  
DR Pfam; PF01001; HCV\_NS4b; 1.  
DR Pfam; PF01506; HCV\_NS5a; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR Pfam; PF00998; Viral\_RdRP; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.

SQ	SEQUENCE	3010 AA; 326838 MW; 58E3BD4140B588 CRC64;	
	Query Match	97.6%; Score 10218; DB 2; Length 3010;	
	Best Local Similarity	96.7%; Pred. No. 0;	
	Matches 1920; Conservative	39; Mismatches 26; Indels 0; Gaps 0;	
QY	1	MAPTAYSOQTRGLGCIITSLGRDNQVEGEVQVNSTAQSELATCVNGVCWTVYHGA	60
DB	1026	LAPITAYSOQTRGLGCIITSLGRDNQVEGEVQVNSTAQSELATCVNGVCWTVYHGA	1085
QY	61	GSKTLAGKPGPIITQMYINVQDQVLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR	120
DB	1086	GSKTLAGKPGPIITQMYINVQDQVLVGMQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR	1145
QY	121	GDGRGSLSPRPVSYLKGSSGGPILLCPSPGHAVGIFPRAAVCTRGVAKAVDFVVPVSMETTM	180
DB	1146	GDGRGSLSPRPISYLYLKGSSGGPILLCPSPGHAVGIFPRAAVCTRGVAKAVDFVVPVSMETTM	1205
QY	181	RSPVFTDNSSPPAVPOTFOVAHLHAPTSGSKSTKVPAAQAQYKVLAVNPSVAATLFGF	240
DB	1206	RSPVFTDNSSPPAVPOTFOVAHLHAPTSGSKSTKVPAAQAQYKVLAVNPSVAATLFGF	1265
QY	241	AYMSKAHGIDENIRTVRTITTCGAPITYSYTKFLADGGCGGAYDIIICDECHSTDSTT	300
DB	1266	AYMSKAHGIDENIRTVRTITTCGAPITYSYTKFLADGGCGGAYDIIICDECHSTDSTT	1325
QY	301	ILGIGTVLDAQETAGARLVVLAATPPGCVTVPHNPINEEVALSGBIPYKAIPIETI	360
DB	1326	ILGIGTVLDAQETAGARLVVLAATPPGCVTVPHNPINEEVALSGBIPYKAIPIETI	1385
QY	361	KGRHLIFCHSKKCDLAAKLGLGNVAVYRGLDVSIVPTSGDVIIVATDALMTGFT	420
DB	1386	KGRHLIFCHSKKCDLAAKLGLGNVAVYRGLDVSIVPTSGDVIIVATDALMTGFT	1445
QY	421	GFDSVIDNCNTVQTQVDFSLDPTFTTETTPQDAVRSQRRGRTGRGMGIYRFVTPG	480
DB	1446	GFDSVIDNCNTVQTQVDFSLDPTFTTETTPQDAVRSQRRGRTGRGMGIYRFVTPG	1505
QY	481	ERSGMPFDSVLCEDYDAGAWBELTAPETSRLRAYLNTPLVPCQDHLFEWESVPTGL	540
DB	1506	ERSGMPFDSVLCEDYDAGAWBELTAPETSRLRAYLNTPLVPCQDHLFEWESVPTGL	1565
QY	541	THIDAFSLQTKAGDNFPVLYAVQATVCARAQAPPSWDMMKCLIRLKTPLHGPPLL	600
DB	1566	THIDAFSLQTKAGDNFPVLYAVQATVCARAQAPPSWDMMKCLIRLKTPLHGPPLL	1625
QY	601	YRLGAVQNEVTTTHPTIKYIMACMSADLEVYVSTWLVGGVLAALAAAYCLITTSQSVTVGR	660
DB	1626	YRLGAVQNEVTTTHPTIKYIMACMSADLEVYVSTWLVGGVLAALAAAYCLITTSQSVTVGR	1685
QY	661	IILSGKPAIIPDREVLRYEFDEMEECASHLPYIEQGMQLAEQKQKAIQLLOTATKQAEA	720
DB	1686	IILSGKPAIIPDREVLRYEFDEMEECASHLPYIEQGMQLAEQKQKAIQLLOTATKQAEA	1745
QY	721	AAPVSKWRTLEAFWAKHWNFTSGIYLAGLSTLPGNPAIASLMAFTASITSPLTQH	780
DB	1746	AAPVSKWRTLEAFWAKHWNFTSGIYLAGLSTLPGNPAIASLMAFTASITSPLTQH	1805
QY	781	TLLFNILGWWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYAGVAGALVA	840
DB	1806	TLLFNILGWWAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYAGVAGALVA	1865
QY	841	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAFASR	900
DB	1866	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAFASR	1925
QY	901	GNHVSPTHYVPESDAAARVQIILSLTITQLLKELHOMINEDCSTPCSGSLRDVMDWIC	960
DB	1926	GNHVSPTHYVPESDAAARVQIILSLTITQLLKELHOMINEDCSTPCSGSLRDVMDWIC	1985
QY	961	TVLTDFKTLQSKLLPRLPGVPPFPFCQYKGYVWRGDGIMQTTCPGCAQITGHVXGSMR	1020
DB	1986	TVLTDFKTLQSKLLPRLPGVPPFPFCQYKGYVWRGDGIMQTTCPGCAQITGHVXGSMR	2045

QY	1021	IVGPRTCSNTHGTFPPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDPHYVTGM	1080
DB	2046	IVGPRTCSNTHGTFPPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDPHYVTGM	2105
QY	1081	TTDNVAKPCQVPAPEPTEVDGVLRLHRYAPACKPLAREEYVTLVGLNOYVGSOLPEPE	1140
DB	2106	TTDNVAKPCQVPAPEPTEVDGVLRLHRYAPACKPLAREEYVTLVGLNOYVGSOLPEPE	2165
QY	1141	PDVAVLTSMLTDFSHITAEAKERLARGSPSPSLASSASQSLAPSLKATCTTRHDSPAD	1200
DB	2166	PDVAVLTSMLTDFSHITAEAKERLARGSPSPSLASSASQSLAPSLKATCTTRHDSPAD	2225
QY	1201	LIENALLWRQEMGNITRVSEKENVILDSFEPLOABEDREVSVPABILLRRSKFFPRAM	1260
DB	2226	LIENALLWRQEMGNITRVSEKENVILDSFEPLOABEDREVSVPABILLRRSKFFPAL	2285
QY	1261	PIWARPDPNPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA	1320
DB	2286	PWARPDPNPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA	2345
QY	1321	ELATKTFGSSSESAVDSGTATSPDQPSDDGDAGSDVESYSSMPPLEGEPCDPLDSGWS	1380
DB	2346	ELATKTFGSSSESAVDSGTATSPDQPSDDGDAGSDVESYSSMPPLEGEPCDPLDSGWS	2405
QY	1381	STVSEBESASEDVCCSMTYWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS	1440
DB	2406	STVSEBESASEDVCCSMTYWTGALITPCAABESKLPINALSNSLLRHHNLVYATTSRAS	2465
QY	1441	LROKVTDFRLOVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK	1500
DB	2466	OROKVTDFRLOVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK	2525
QY	1501	DVNLSKAVNHHSVWKOLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG	1560
DB	2526	DVNLSKAVNHHSVWKOLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG	2585
QY	1561	VRVCEKMAVLDVYSTIPOAVMGSSYGFQVSPGQVRFVFNAAKAKCPMGFAYDTRCFDS	1620
DB	2586	VRVCEKMAVLDVYSTIPOAVMGSSYGFQVSPGQVRFVFNAAKAKCPMGFAYDTRCFDS	2645
QY	1621	TVTENDIRVEESYQCCDLAPARQALRSILTERLYIGGPLTNSKGNCGYRRCRAGVLT	1680
DB	2646	TVTENDIRVEESYQCCDLAPARQALRSILTERLYIGGPLTNSKGNCGYRRCRAGVLT	2705
QY	1681	TSCGNTLTCYKAAAAACRAAKLODCTMLVCGDDLVIICESAGTQEDAEASLRATEMTRY	1740
DB	2706	TSCGNTLTCYKAAAAACRAAKLODCTMLVCGDDLVIICESAGTQEDAEASLRATEMTRY	2765
QY	1741	SAPPGPDPPEYDLELITSCSSNVVAHDASGRVYVLTDRDPTPLARAANETARHTPVN	1800
DB	2766	SAPPGPDPPEYDLELITSCSSNVVAHDASGRVYVLTDRDPTPLARAANETARHTPVN	2825
QY	1801	SWLGNIMVAPTLMWARMILMTHFFSILLAQEULEKALDCOYGCACYSEPLDLPQIIQRL	1860
DB	2826	SWLGNIMVAPTLMWARMILMTHFFSILLAQEULEKALDCOYGCACYSEPLDLPQIIQRL	2885
QY	1861	HGLSAPLSHSYSPGEINRVASCLRLKGVPLRVWRHARSVRARLLSQGGRATCGKLYF	1920
DB	2886	HGLSAPLSHSYSPGEINRVASCLRLKGVPLRVWRHARSVRARLLSQGGRATCGKLYF	2945
QY	1921	NWAVRTKLTPTIPAAASQDLSSVFAVYSGGDIYHLSRARPRWFWMLLLLSVGVGIY	1980
DB	2946	NWAVRTKLTPTIPAAASQDLSSVFAVYSGGDIYHLSRARPRWFWMLLLLSVGVGIY	3005
QY	1981	LLPNR 1985	
DB	3006	LLPNR 3010	

RESULT 8  
Q7T4V8  
ID Q7T4V8 PRELIMINARY; PRT; 1984 AA.

```
AC Q7TAV8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NS protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2694192; PubMed=12810084;
RA Kato N., Sugiyama K., Namba K., Dansako H., Nakamura T., Takami M.,
RA Naka K., Nozaki A., Shintouo K.;
RT "Establishment of a hepatitis C virus subgenomic replicon derived from
RL human hepatocytes infected in vitro.";
RL Biochem. Biophys. Res. Commun. 306:756-766(2003).
DR EMBL; AB109543; BAC77767.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0013079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cyt_c_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
FT NON_TER 1 1
SQ SEQUENCE 1984 AA; 214359 MW; 8AA8198D2C7B291C CRC64;

Query Match 97.5%; Score 10206; DB 2; Length 1984;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1921; Conservative 35; Mismatches 28; Indels 0; Gaps 0;

QY 2 APITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVYSTATQSFATPCVNGCMTVYHGAG 61
Db 1 APITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVYSTATQSFATPCVNGCMTVYHGAG 60
QY 62 SKTLAPKGPITOMYTNVDVLGVCAPEGASLTCTCGSDLYLVTHADVIPIVRRG 121
Db 61 SKTLAPKGPITOMYTNVDVLGVCAPEGASLTCTCGSDLYLVTHADVIPIVRRG 120
QY 122 DSRGSLSPRPVSYLKGSSGGPILCPSGHAGVIFRAAVCTRGVAKAVDFVPVSEMETTWR 181
Db 121 DNRGSLSPRPVSYLKGSSGGPILCPSGHAGVIFRAAVCTRGVAKAVDFVPVSEMETTWR 180
QY 182 SPVFTDNSSPPAVPTQFVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLGFCA 241
Db 181 SPVFTDNSSPPAVPTQFVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLGFCA 240
QY 242 YMSKAGIDPNRTGVRTITTCAPITYSTYKGLADGCGGAYDIIICDECHSTDSTTI 301
Db 241 YMSKAGIDPNRTGVRTITTCAPITYSTYKGLADGCGGAYDIIICDECHSTDSTSI 300
QY 302 LGIGTVLDQAGTAGARLVLTATPPGSGTVVPHNPNEEVALSNTGEIPYGKAFPLEAK 361
Db 301 LGIGTVLDQAGTAGARLVLTATPPGSGTVVPHNPNEEVALSNTGEIPYGKAFPLEAK 360
QY 362 GGRHLIFCHSKKKCDLAELAKISGLGINAVAYRGLDVSVIPSGDVIVVATDALMTGFTG 421
Db 361 GGRHLIFCHSKKKCDLAELAKISGLGINAVAYRGLDVSVIPSGDVIVVATDALMTGFTG 420
QY 422 DFDSVIDCNTCVTQTVDSLDFTFIETTTVPQDAVSRQSQRGTRGRGMGYRFRVTGPE 481
Db 421 DFDSVIDCNTCVTQTVDSLDFTFIETTTVPQDAVSRQSQRGTRGRGMGYRFRVTGPE 480
QY 482 RPSGMFDSVLCCEYDAGCAWYELTPAETSVRRLAYLNTPLPVCQDHLFEFVESFTGTL 541
Db 481 RPSGMFDSVLCCEYDAGCAWYELTPAETSVRRLAYLNTPLPVCQDHLFEFVESFTGTL 540
QY 542 HIDAHLFSLQTKOAGDNFPYLVAYQATVCARAQAPPSQDMWKKLIRLKPTHLHGTPPLY 601
Db 541 HIDAHLFSLQTKOAGDNFPYLVAYQATVCARAQAPPSQDMWKKLIRLKPTHLHGTPPLY 600
QY 602 RLGAQNVEVTTTHPIITKTFIMACMSADLEVTSTWLVGVLAALAAAYCLTTCGSVIVGRI 661
Db 601 RLGAQNVEVTTTHPIITKTFIMACMSADLEVTSTWLVGVLAALAAAYCLTTCGSVIVGRI 660
QY 662 ILSGKPAIIPDREVILYREFDEMEECASHLPYIEQGMQLAEQFKQALGLQATKQAEAA 721
Db 661 ILSGKPAIIPDREVILYREFDEMEECASHLPYIEQGMQLAEQFKQALGLQATKQAEAA 720
QY 722 APVSKWRTLEAFWAKHMMNFIISQIYLAGLSTLPGNPALIASLMAFTSIISPLTQHT 781
Db 721 APVSKWRTLEAFWAKHMMNFIISQIYLAGLSTLPGNPALIASLMAFTSIISPLTQHT 780
QY 782 LLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLGKLVLDVILLAGVAGVAGALVAF 841
Db 781 LLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLGKLVLDVILLAGVAGVAGALVAF 840
QY 842 KVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVMMNRLIAPASRG 901
Db 841 KVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVMMNRLIAPASRG 900
QY 902 NHVSPTHYVPESDAARVTQILSSITITQLLKRLLQWINEDECTPCSGSLWLDVMDWICT 961
Db 901 NHVSPTHYVPESDAARVTQILSSITITQLLKRLLQWINEDECTPCSGSLWLDVMDWICT 960
QY 962 VLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDGMOTTCPCGAQITGHVKGSMRI 1021
Db 961 VLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDGMOTTCPCGAQITGHVKGSMRI 1020
QY 1022 VGPRTCSNTWHTGTFPINAAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPHYVTGMT 1081
Db 1021 VGPRTCSNTWHTGTFPINAAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPHYVTGMT 1080
QY 1082 TQNVKPCQCPVAPPEFTEVDGVRHLHYAPACKPLREBVTFLVGLNQYLVGSQLPCEPEP 1141
Db 1081 TQNVKPCQCPVAPPEFTEVDGVRHLHYAPACKPLREBVTFLVGLNQYLVGSQLPCEPEP 1140
QY 1142 DYAVLTSLMTDPSHTTAETAKRRLARGSPPLASSASOLSAFSLKATCTTRHSDPADL 1201
Db 1141 DTVLTSLMTDPSHTTAETAKRRLARGSPPLASSASOLSAFSLKATCTTRHSDPADL 1200
QY 1202 IEANLLWRQEMGNGITRVESENKVVILDSFELQAEDEDEREVSFAELIRSRKPPRAMP 1261
Db 1201 IEANLLWRQEMGNGITRVESENKVVILDSFELQAEDEDEREVSFAELIRSRKPPRAMP 1260
QY 1262 IWARPDYNPPLLESNKDDYVPPVHVGCPPLPAKAPPITPPRRKRTVILSESTVSALAE 1321
Db 1261 IWARPDYNPPLLESNKDDYVPPVHVGCPPLPAKAPPITPPRRKRTVILSESTVSALAE 1320
QY 1322 LATKTFGSSSESAVDSGTATASPDOPSDDGADAGSDVESYSSMPLEGEPCDPLSDGWS 1381
Db 1321 LATKTFGSSSESAVDSGTATASPDOPSDDGADAGSDVESYSSMPLEGEPCDPLSDGWS 1380
QY 1382 TVSEBSEBVDCCSSYWTGALITPCAAEETKLPINALSNLLRHHNLVYATTSSASL 1441
Db 1381 TVSEBSEBVDCCSSYWTGALITPCAAEETKLPINALSNLLRHHNLVYATTSSASL 1440
```

Db 1381 TVSEASEDVVCCSMYSYTWTCALITPCAAEESKLPINALSNSLLRHNNVYATTSSASQ 1440

QY 1442 RQKVTEDRLVDHDDHYEDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKD 1501

Db 1441 RQKVTEDRLVDHDDHYEDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKD 1500

QY 1502 VRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVCVQPEKGRKPABLIIVFPDLGV 1561

Db 1501 VRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVCVQPEKGRKPABLIIVFPDLGV 1560

QY 1562 RVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDIRCEPDS 1621

Db 1561 RVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDIRCEPDS 1620

QY 1622 VTENDIRVEESIYCCDLAPARQAIRSLTERLYIGGPLTNSKQNCYRCRASGVLT 1681

Db 1621 VTENDIRVEESIYCCDLAPARQAIRSLTERLYIGGPLTNSKQNCYRCRASGVLT 1680

QY 1682 SCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1741

Db 1681 SCGNLTLCYLKASACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740

QY 1742 APQDPPKPEYDLBELITSCSNVSAHDASGRVYVLTTRDPTTFLARAAMETARHTPVNS 1801

Db 1741 APQDPPKPEYDLBELITSCSNVSAHDASGRVYVLTTRDPTTFLARAAMETARHTPVNS 1800

QY 1802 WLGNIMVAPTLWARMTLMTFFSILLAOEQLKALDCQIYGATYIEPDLPLQIIQRLH 1861

Db 1801 WLGNIMVAPTLWARMTLMTFFSILLAOEQLKALDCQIYGATYIEPDLPLQIIQRLH 1860

QY 1862 GLSAFSLHSYSPGEINRVASCLRLKGLVPLRVWHRARSVRARLLSOGGAAATCGKYLEN 1921

Db 1861 GLSAFSLHSYSPGEINRVASCLRLKGLVPLRVWHRARSVRARLLSOGGAAATCGKYLEN 1920

QY 1922 WAVRTKLKLTPIPAASQDLSSVNFVAGYSGGDIYVHLSRARPRFWMWCLLLSVGVGIYL 1981

Db 1921 WAVRTKLKLTPIPAASQDLSSVNFVAGYSGGDIYVHLSRARPRFWMWCLLLSVGVGIYL 1980

QY 1982 LPNR 1985

Db 1981 LPNR 1984

RESULT 9

P90194

ID P90194 PRELIMINARY; PRT; 3010 AA.

AC P90194;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)

DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Polyprotein.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI\_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HCV-1b;

RX MEDLINE=95340824; PubMed=7542279;

RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T., Yamamoto C., Izumi N., Marumo F., Sato C.;

RT "Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b";

RL J. Clin. Invest. 96:224-230 (1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=HCV-1b;

RA Enomoto N.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; D50481; BAA09072.1; -.

DR FIR; A61196; A61196.

DR FIR; PQ0246; PQ0246.

DR FIR; PQ0254; PQ0254.

PIR; PS0329; PS0329.

HSP; P26663; LUXP.

GO; GO:0016021; C: integral to membrane; IEA.

GO; GO:0019028; C: viral capsid; IEA.

GO; GO:0019031; C: viral envelope; IEA.

GO; GO:0005524; F: ATP binding; IEA.

GO; GO:0008026; F: ATP-dependent helicase activity; IEA.

GO; GO:0003723; F: RNA binding; IEA.

GO; GO:0003723; F: RNA binding; IEA.

GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.

GO; GO:0005198; F: serine-type peptidase activity; IEA.

GO; GO:0006508; F: structural molecule activity; IEA.

GO; GO:0006508; F: proteolysis and peptidolysis; IEA.

GO; GO:0006350; P: transcription; IEA.

GO; GO:0019079; P: viral genome replication; IEA.

GO; GO:0019087; P: viral transformation; IEA.

InterPro; IPR000345; CytC\_heme\_BS.

InterPro; IPR001410; DEAD.

InterPro; IPR002522; HCV\_capsid.

InterPro; IPR002521; HCV\_core.

InterPro; IPR002519; HCV\_env.

InterPro; IPR002531; HCV\_NS1.

InterPro; IPR000745; HCV\_NS4a.

InterPro; IPR001490; HCV\_NS4b.

InterPro; IPR002868; HCV\_NS5a.

InterPro; IPR002166; HCV\_RdRP.

InterPro; IPR001650; Helicase\_C.

InterPro; IPR004109; Peptidase\_S29.

InterPro; IPR009003; Peptidase\_S29.

InterPro; IPR002518; Pept\_U39\_HCV\_NS2.

InterPro; IPR007095; RNA\_pol\_DS\_PS.

InterPro; IPR007094; RNA\_pol\_PSVir.

Pfam; PF01543; HCV\_capsid; 1.

Pfam; PF01542; HCV\_core; 1.

Pfam; PF01539; HCV\_env; 1.

Pfam; PF01560; HCV\_NS1; 1.

Pfam; PF01538; HCV\_NS2; 1.

Pfam; PF02907; HCV\_NS3; 1.

Pfam; PF01006; HCV\_NS4a; 1.

Pfam; PF01001; HCV\_NS4b; 1.

Pfam; PF01506; HCV\_NS5a; 1.

Pfam; PF00271; Helicase\_C; 1.

Pfam; PF00998; Viral\_RdRP; 1.

SMART; SM00487; DEXDC; 1.

PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.

Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

Polyprotein; Transmembrane.

CHAIN 810 1026 NS2.

CHAIN 1027 1657 NS3.

CHAIN 1658 1711 NS4a.

CHAIN 1 191 core protein.

CHAIN 1712 1972 NS4b.

CHAIN 1973 2419 NS5a.

CHAIN 2420 3010 NS5b.

CHAIN 192 383 E1.

CHAIN 384 809 E2.

SEQUENCE 3010 AA; 326816 MW; 98D5C2A2D47FD011 CRC64;

Query Match 97.5%; Score 10199; DB 2; Length 3010;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 1920; Conservative 34; Mismatches 31; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCIITSLTGRDNRQVEGEVQVWSTATQSFATCNGVCWTYHGA 60

1026 LAPITAYSQOTRGLGCIITSLTGRDNRQVEGEVQVWSTATQSFATCNGVCWTYHGA 1085

QY 61 GSKTLAGPKGPITQMTYNTVDQLVQWAPPGARSLTPTCGSSDLYLVTRHADVIPVRR 120

1086 GSKTLAGPKGPITQMTYNTVDQLVQWAPPGARSLTPTCGSSDLYLVTRHADVIPVRR 1145

QY 121 GDSRGLSLSPRVSYLKGSSGGLPLCPSPGHVGIIFRAAVCTRGVAKAVDFVPSMETTM 180

1146 GDSRGLSLSPRVSYLKGSSGGLPLCPSPGHVGIIFRAAVCTRGVAKAVDFVPSMETTM 1205

191 RSPVPTDSSPPAVPQTTOVAHLHAPTGSKSTKVPAAAYAGYKVLVNLPSVAATLGGF 240  
1206 RSPVPTDSSPPAVPQTTOVAHLHAPTGSKSNKVPVEYAAQGYKVLVNLPSVAATLGGF 1265  
241 AYMSKAHGIDPNIRTVGRTITTGAPITYSTYKFLADGCGSGAYDIIICDCHSTDSST 300  
1266 AYMSKAHGVDNIRTVGRTITTGAPITYSTYKFLADGCGSGAYDIIICDCHSTDSST 1325  
301 ILIGITVLDQAETAGARLVLTATATPPGSVTVPHENIEVALSSTGEIPFYGKAIPETI 360  
1326 ILIGITVLDQAETAGARLVLTATATPPGSVTVPHENIEVALSSTGEIPFYGKAIPETI 1385  
361 KGRHLLIFCHSKKKCDELAAKLSGLGNAYVYRGLDVSVIPTSDDVIWATDALMTGFT 420  
1386 KGRHLLIFCHSKKKCDELAAKLSGLGNAYVYRGLDVSVIPASGDVVVWATDALMTGFT 1445  
421 GPDFSVIDNCTVOTVDFSLDPTFTIETTVPDQAVSRQRRGRTGRGMGIYRFVTPG 480  
1446 GPDFSVIDNCTVOTVDFSLDPTFTIETTVPDQAVSRQRRGRTGRGAGIYRFVTPG 1505  
481 ERPSGFDSSVLCCEYDAGCAYVELTPAETSVRLRAYLNTPLGVPQDHLFEWESVFTGL 540  
1506 ERPSGFDSSVLCCEYDAGCAYVELTPAETSVRLRAYLNTPLGVPQDHLFEWESVFTGL 1565  
541 THIDAHFLSOTKQAGNFPYLYAYQATVCARAQAPPSWDQWKCLIRLKPTLHGPTLL 600  
1566 THIDAHFLSOTKQAGNFPYLYAYQATVCARAQAPPSWDQWKCLIRLKPTLHGPTLL 1625  
601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVTSTWLVGVLAALAAAYCLITGSSVIVGR 660  
1626 YRLGAVQNEVTTTHPTIKYIMACMSADLEVTSTWLVGVLAALAAAYCLITGSSVIVGR 1685  
661 IILSGKPAIIPDREVLRYREFDEMEECASHLYPIEQMGQALABOFKQKALGLQATKQAEA 720  
1686 IILSGKPAIIPDREVLRYREFDEMEECASHLYPIEQMGQALABOFKQKALGLQATKQAEA 1745  
721 AAPVSKWRTLEAFWAKHWNFIISGIVYLAGLSTLPGNPATASLMAFTASITSPUTTOH 780  
1746 AAPVSKWRTLEAFWAKHWNFIISGIVYLAGLSTLPGNPATASLMAFTASITSPUTTOH 1805  
781 TLLFNILGCVAAQALAPPASAAAFVAGIAGAAVSGIGLGVKLVLDILAGYAGVAGALVA 840  
1806 TLLFNILGCVAAQALAPPASAAAFVAGIAGAAVSGIGLGVKLVLDILAGYAGVAGALVA 1865  
841 FKVMGEMPTEDLVNMLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAPASR 900  
1866 FKVMGEMPTEDLVNMLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAPASR 1925  
901 GNVHSPTHVYVPESDAARVTOILSLTITQLKRLHWINEDCSTPCSGSWLRDVMWDIC 960  
1926 GNVHSPTHVYVPESDAARVTOILSLTITQLKRLHWINEDCSTPCSGSWLRDVMWDIC 1985  
961 TVLTDFKTLQSKLLPRLPGVPPFCQYGYKVMRGDGMQTTCPGCAQITGHVKNGSMR 1020  
1986 TVLSDFKTLQSKRLPRLPGVPLSCQYGYKVMRGDGMQTTCPGCAQITGHVKNGSMR 2045  
1021 IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALNRVAABEYVEVTVRGDFHYVTGM 1080  
2046 IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALNRVAABEYVEVTVRGDFHYVTGM 2105  
1081 TTDNVKCPQVPAPEPTEVDGRLHRYAPACKPLLRREYVTVLVLNOYLVSQLPCEPE 1140  
2106 TTDNLKCPQVPAPEPTEVDGRLHRYAPACKPLLRREYVTVLVLNOYLVSQLPCEPE 2165  
1141 PDVAVLTSMLTDFSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTHRDSPAD 1200  
2166 PDVAVLTSMLTDFSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTHRDSPAD 2225  
1201 LIEANLLWRQMGNNITRVSENNKVILDSFEPLQAEEDEREVSVAEIIILRRSKKPPRAM 1260  
2226 LIEANLLWRQMGNNITRVSENNKVILDSFEPLQAEEDEREVSVAEIIILRRSKKPPRAM 2285  
1261 PIWARPDPYNPLESWKDPDVPVPHVHCPLPPAKAPPIPPRRKRTVWLSESTVSSALA 1320

2286 PIWARPDPYNPLESWKDPDVPVPHVHCPLPPAKAPPIPPRRKRTVWLSESTVSSALA 2345  
1321 ELATKTFGSSSAVDSCGTATASPDQDDGAGSDVESYSMPLEGEPPDLSGDSW 1380  
2346 ELATKTFGSSSAVDSCGTATASPDQDDGAGSDVESYSMPLEGEPPDLSGDSW 2405  
1381 STVSEAEASDVVCCSMSTWTGALITPCAABETKLPINALSNLRRHNLVATTSRAS 1440  
2406 STVSEAEASDVVCCSMSTWTGALITPCAABESKLPINALSNLRRHNLVATTSRAS 2465  
1441 LRQKVTEDRLOVLDDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 1500  
2466 QRQKVTEDRLOVLDDHYRDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 2525  
1501 DVNRLSSKAVNHIRSVWKOLLEDTETPTDITIMAKNEVFCVQPEKGRKPAELIIFPDIG 1560  
2526 DVNRLSSKAVNHIRSVWKOLLEDTETPTDITIMAKNEVFCVQPEKGRKPAELIIFPDIG 2585  
1561 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
2586 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 2645  
1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNCYRRCRAGSVLT 1680  
2646 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNCYRRCRAGSVLT 2705  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDEASIRAFTEAMTRY 1740  
2706 TSCGNLTLCYLKATAACRAAKLQDCTMLVCGDDLVVICESAGTOEDEASIRAFTEAMTRY 2765  
1741 SAPPGDPKPEYDLELITSCSSNVSVAHDSKRVVYLTRDPTPLARAWEETARTHPVN 1800  
2766 SAPPGDPKPEYDLELITSCSSNVSVAHDSKRVVYLTRDPTPLARAWEETARTHPVN 2825  
1801 SWLGNIIIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 1860  
2826 SWLGNIIIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 2885  
1861 HGLSAPLSHSYSGEINRVASCLKLGVPPLVWRHRSVRARLLSOGGRAATCGKLYF 1920  
2886 HGLSAPLSHSYSGEINRVASCLKLGVPPLVWRHRSVRARLLSOGGRAATCGKLYF 2945  
1921 NNAVRTKLLTPIPAASQDLSSWFSVAGYSGGDIHSLSRARPRFWFWCILLISVGVGIY 1980  
2946 NNAVRTKLLTPIPAASQDLSSWFSVAGYSGGDIHSLSRARPRFWFWCILLISVGVGIY 3005  
1981 LLPNR 1985  
3006 LLPNR 3010

## RESULT 10

Q99AU2 PRELIMINARY; PRT: 3010 AA.  
AC Q99AU2;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus type 1b.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 1.  
OX NCBI\_TaxID=31647;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=chimera of HCV-BK;  
RX MEDLINE=21534507; PubMed=11677216;  
RA Thomson M., Nascimbeni M., Gonzales S., Murthy K.K., Rehmann B.,  
Liang T.J.;  
RT "Emergence of a distinct pattern of viral mutations in chimpanzees  
infected with a homogeneous inoculum of hepatitis C virus.";  
RL Gastroenterology 121:1226-1233(2001).





1381 STVSEASEDDVCCSMYSWTGALITPCAAREBTKLPINALSNLLRHHNVATTSSRAS 1440  
2406 STVSEASEDDVCCSMYSWTGALITPCAAREBTKLPINALSNLLRHHNVATTSSRASG 2465  
1441 LQKQVTFDLQVLDHVDVLEKEMKAKASTYKAKLLSVEEACKLTPPHSARSKEGYGAK 1500  
2466 LQKQVTFDLQVLDHVDVLEKEMKAKASTYKAKLLSVEEACKLTPPHSARSKEGYGAK 2525  
1501 DYNRLSSKAVNHRSWKDLLEDTETPTIDTTIMAKNEVFCVQPEKGGKRPABLIIVPDLG 1560  
2526 DYNRLSSKAVNHRSWKDLLEDTETPTIDTTIMAKNEVFCVQPEKGGKRPABLIIVPDLG 2585  
1561 VVACEKMALYDVVSTLPQAVMGSSYGFQYSPQGVFEFLVNAWKAKCPEMFAVDYRCRPS 1620  
2586 VVACEKMALYDVVSTLPQAVMGSSYGFQYSPQGVFEFLVNTWKSXKPMGFYDRCRPS 2645  
1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGELTNSKQNGCYRRCRASGVL 1680  
2646 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGELTNSKQNGCYRRCRASGVL 2705  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDASLRAFTAMTRY 1740  
2706 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDASLRAFTAMTRY 2765  
1741 SAPPDGPPEYDLELITSCSNVSVAHDAKGRVYVLTTRDPTTPLARAAMETARHTPVN 1800  
2766 SAPPDGPPEYDLELITSCSNVSVAHDAKGRVYVLTTRDPTTPLARAAMETARHTPVN 2825  
1801 SWLGNIMYAPTLMWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIQL 1860  
2826 SWLGNIMYAPTLMWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIQL 2885  
1861 HGLSFAFSLHSYSPGEINRVASCLKLGVPPLVWHRARSVPARLLSQGGRATCGKLYF 1920  
2886 HGLSFAFSLHSYSPGEINRVASCLKLGVPPLVWHRARSVPARLLSQGGRATCGKLYF 2945  
1921 NNAVATKILKLTIPAAQDLDSWFWAGYSGGDIYHLSRARPFRWFMCLLLLSVGVGTY 1980  
2946 NNAVATKILKLTIPAAQDLDSWFWAGYSGGDIYHLSRARPFRWFMCLLLLSVGVGTY 3005  
1981 LLPNR 1985  
3006 LLPNR 3010

RESULT 11  
Q9DTE9  
ID Q9DTE9 PRELIMINARY; PRT: 3010 AA.  
AC Q9DTE9;  
DT 01-MAR-2001 (TremBLrel. 16, Created)  
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]\_TaxID=11103;  
RP SEQUENCE FROM N.A.  
RC TISSUE=Serum;  
RA Hatahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,  
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,  
RA Mishiro S.;  
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
RT with hepatocellular carcinoma: the 'progression score' revisited."  
RL Hepatol. Res. 20:161-171(2001).  
DR EMBL; AB049088; EAB18801.1; -;  
DR PIR; A61196; A61196.  
DR PIR; PQ0246; PQ0246.  
DR PIR; PQ0804; PQ0804.  
DR PIR; PS0329; PS0329.  
DR HSSP; P26663; LJPX.  
DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; F:transcription; IEA.  
DR GO; GO:0019079; F:viral genome replication; IEA.  
DR GO; GO:0019087; F:viral transformation; IEA.  
DR InterPro; IPR00345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD\_capsid.  
DR InterPro; IPR002522; HCV\_capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR InterPro; IPR00745; HCV NS4a.  
DR InterPro; IPR001490; HCV NS4b.  
DR InterPro; IPR002868; HCV NS5a.  
DR InterPro; IPR002166; HCV RdRp.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR004109; Peptidase\_S29.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
DR Pfam; PF01538; HCV NS2; 1.  
DR Pfam; PF02907; HCV NS3; 1.  
DR Pfam; PF01006; HCV NS4a; 1.  
DR Pfam; PF01001; HCV NS4b; 1.  
DR Pfam; PF01506; HCV NS5a; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR Pfam; PF00998; Viral\_RdRp; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
SQ SEQUENCE 3010 AA; 326780 MW; 668CFFEA5FEC3658 CRC64;  
Query Match 97.3%; Score 10187; DB 2; Length 3010;  
Best Local Similarity 96.5%; Pred.No. 0;  
Matches 1915; Conservative 41; Mismatches 29; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVWSTATQSFILATCVNGVCTVYHGA 60  
Db 1026 LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVWSTATQSFILATCVNGVCTVYHGA 1085  
QY 61 GSKTLAGPKGITQMYTNDQDLVQWQAPPAGARSITPCTCGSSDLYLVTRHADVPVRRR 120  
Db 1086 GSKTLAGPKGITQMYTNDQDLVQWQAPPAGARSITPCTCGSSDLYLVTRHADVPVRRR 1145  
QY 121 GDSRGLSPRPVSYLKGSSGGLPCPSGHAVGIFFRAAVCTRGVAKAVDFVVESEMETTM 180  
Db 1146 GDSRGLSPRPVSYLKGSSGGLPCPSGHAVGIFFRAAVCTRGVAKAVDFVVESEMETTM 1205  
QY 181 RSPVFTDNSSPPAVPOTQVAHLHAPTQSGKSTKVPAAAYAAQYKVLVNPVVAATLPGF 240  
Db 1206 RSPVFTDNSSPPAVPOTQVAHLHAPTQSGKSTKVPAAAYAAQYKVLVNPVVAATLPGF 1265  
QY 241 AYMSKAHGIDNPRTGVRTITTTGAPITYSTYTGKFLADGCGSGGAYDIIICDCHSTDSTT 300  
Db 1266 AYMSKAHGIDNPRTGVRTITTTGAPITYSTYTGKFLADGCGSGGAYDIIICDCHSTDSTT 1325  
QY 301 ILGIGTVLDQARTAGARLVVLATATPPGCVTVPHNIEEVALSSTGEIIPFYKAIPIETI 360  
Db 1326 ILGIGTVLDQARTAGARLVVLATATPPGCVTVPHNIEEVALSSTGEIIPFYKAIPIETI 1385

361 KGRHLLFCHSKKCDLAAKLSGLGNNAVAYRGLDVSIVPTSDVIVVATDALMTGFT 420  
1386 KGRHLLFCHSKKCDLAAKLSGLGNNAVAYRGLDVSIVPTSDVIVVATDALMTGFT 1445  
421 GDFSDVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVSRORRGRGTGRGMGIYRFVTPG 480  
1446 GDFSDVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVSRORRGRGTGRGMGIYRFVTPG 1505  
481 ERPSGMEFSSVLCBYDAGCAWYBELTPTAETSVRLRAYINTPGLPVCQDHLFEWESVFTGL 540  
1506 ERPSGMEFSSVLCBYDAGCAWYBELTPTAETSVRLRAYINTPGLPVCQDHLFEWESVFTGL 1565  
541 THIDAHLFSLQKAGDNPFYLVAYQATVCARAQAPPSWQMKCLIRLXPTLHGPTPL 600  
1566 THIDAHLFSLQKAGDNPFYLVAYQATVCARAQAPPSWQMKCLIRLXPTLHGPTPL 1625  
601 YRLGAVQNEVTHPTIKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLTTGWSVIVGR 660  
1626 YRLGAVHNEVTLTHPTIKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLTTGWSVIVGR 1685  
661 IILSGKPAITPDREVLRYREDEBECAHLPYIBQGMOLAEQFKQKAGLQATKQAEA 720  
1686 IILSGKPAITPDREVLRYREDEBECAHLPYIBQGMOLAEQFKQKAGLQATKQAEA 1745  
721 AAPVWESKWTLEAFWAKHWNFTSGIQLAGLSTLPCNPAIASLMAFTASITSPLTQH 780  
1746 AAPVWESKWTLEAFWAKHWNFTSGIQLAGLSTLPCNPAIASLMAFTASITSPLTQH 1805  
781 TLLFNILGGWVAQAAPPSAASAFVAGIAGAAVGSIGLQKVLVDIILAGYAGVAGALVA 840  
1806 TLLFNILGGWVAQAAPPSAASAFVAGIAGAAVGSIGLQKVLVDIILAGYAGVAGALVA 1865  
841 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 900  
1866 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 1925  
901 GNVHSPHYVPESSDAARVTQILSSLTITOLLKELHOMINEDCSTPCSGSMLRDVMDWIC 960  
1926 GNVHSPHYVPESSDAARVTQILSSLTITOLLKELHOMINEDCSTPCSGSMLRDVMDWIC 1985  
961 TVLTDFKTLQSKLLPRLPGVPFFSCQYKGVWRGDGIMQTTCPGCAQITGHVKGSMR 1020  
1986 TVLTDFKTLQSKLLPRLPGVPFFSCQYKGVWRGDGIMQTTCPGCAQISGHVKGSMR 2045  
1021 IVGPRICSNTHGTFPPINAYTTGCTPSPAPNYSRALRWVAABEYVETRVGDFHYVTGM 1080  
2046 IVGPRICSNTHGTFPPINAYTTGCTPSPAPNYSRALRWVAABEYVETRVGDFHYVTGM 2105  
1081 TTDNVKPCQVPAPEFTEVDGVRHLHRYAPACKPLLRREYVTVLGNOLVGSQLPCEPE 1140  
2106 TTDNVKPCQVPAPEFTEVDGVRHLHRYAPACKPLLRREYVTVLGNOLVGSQLPCEPE 2165  
1141 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTTHDSDPAD 1200  
2166 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTTHDSDPAD 2225  
1201 LIENALLRWQEMGNITRVSEKNVILDSFEPLOABEDREVSVPABILRRSRKTFPRAM 1260  
2226 LIENALLRWQEMGNITRVSEKNVILDSFEPLOABEDREVSVPABILRRSRKTFPRAM 2285  
1261 PIWAPDYNPILLESWKDPDYVPPVHGCPLPPAKAPPPIPPRRKRTVLTSTVSSALA 1320  
2286 PIWAPDYNPILLESWKDPDYVPPVHGCPLPPAKAPPPIPPRRKRTVLTSTVSSALA 2345  
1321 ELATKTFGSSSAVDSGTATAPDQSDGAGSDVESYSSMPPLGEFPGDPLSDGSW 1380  
2346 ELATKTFGSSSAVDSGTATAPDQSDGAGSDVESYSSMPPLGEFPGDPLSDGSW 2405  
1381 STVSEBASEDVVCCSMYTWGTALITPCAAEBETKLPINALNSLLRHNNLVYATTSRAS 1440  
2406 STVSEBASEDVVCCSMYTWGTALITPCAAEBETKLPINALNSLLRHNNLVYATTSRAS 2465  
1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVXAKULLSVEEACKLTPPHSARSKFGYGAK 1500

Db 2466 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVXAKULLSVEEACKLTPPHSARSKFGYGAK 2525  
QY 1501 DVRNLSSKANVHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGKPKARLIVFPDLG 1560  
Db 2526 DVRNLSSKANVHRSVWKDLLEDTETPIDTTIMAKNEIFCVQPEKGGKPKARLIVFPDLG 2585  
QY 1561 VRVCEKMAVYDVVSTLPOAVMGSSYGFQYSPGQRFVEELVNAWAKKCPMGFPAYDTRCFDS 1620  
Db 2586 VRVCEKMAVYDVVSTLPOAVMGSSYGFQYSPGQRFVEELVNAWAKKCPMGFPAYDTRCFDS 2645  
QY 1621 TVTENDIRIBESYIQCDDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRASGVLIT 1680  
Db 2646 TVTENDIRIBESYIQCDDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRASGVLIT 2705  
QY 1681 TSCGNTLTCVLAACAAKLODCTMLVCGDDLVVICESAGTOEDEASLRAFTEAMTRY 1740  
Db 2706 TSCGNTLTCVLAACAAKLODCTMLVCGDDLVVICESAGTOEDEASLRAFTEAMTRY 2765  
QY 1741 SAPPGDPKPEYDLELITSCSSNVSVVAHDASGRVYVYLTDRPTTPLARAAMETARHTFPV 1800  
Db 2766 SAPPGDPKPEYDLELITSCSSNVSVVAHDASGRVYVYLTDRPTTPLARAAMETARHTFPV 2825  
QY 1801 SWLGNIIIMYAPTLMWARMILMTHFESSILLAQOLEKALDCOYGCYSIEPLDLPQIIERL 1860  
Db 2826 SWLGNIIIMYAPTLMWARMILMTHFESSILLAQOLEKALDCOYGCYSIEPLDLPQIIERL 2885  
QY 1861 HGLSASFSLHSYSGEINRVASCLRKLVPLRVHRARSVRARLLSQGGRATCGKYL 1920  
Db 2886 HGLSASFSLHSYSGEINRVASCLRKLVPLRVHRARSVRARLLSQGGRATCGKYL 2945  
QY 1921 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHLSRPRPWFMMWCLLLSVGVGIY 1980  
Db 2946 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHLSRPRPWFMMWCLLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

RESULT 12  
Q9QIY8  
ID Q9QIY8 PRELIMINARY; PRT; 3010 AA.  
AC Q9QIY8;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MD2-1;  
RX MEDLINE=20013325; PubMed=10544098;  
RA Nagayama K., Kuroaki M., Enomoto N., Maekawa S.-y., Miyasaka Y.,  
Tazawa J.-i., Izumi N., Marumo F., Sato C.;  
RT "Time-related changes in full-length hepatitis C virus and hepatitis  
activity";  
RL Virology 263:244-253 (1999).  
DR EMBL; AF165047; AAD56182.1; -.  
DR PIR; A61196; A61196.  
DR PIR; P02246; P02246.  
DR PIR; P02251; P02251.  
DR PIR; P02252; P02252.  
DR PIR; P02254; P02254.  
DR HSSP; P26663; IJXP.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.



```
QY 1501 DVRLSSKAVNHRSVWVKMLLEDTEPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560
D 2526 DVRLSSKAVNHRSVWVKMLLEDTEPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585
QY 1561 VRCEKALYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWAKKCPMGFAYDTRCFDS 1620
D 2586 VRCEKALYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWAKKSPMGFAYDTRCFDS 2645
QY 1621 TVTENDIRVEESIYCCDLAPEARQAIISLTERLYIGGLPLNSKQNGCYRRCRASGVLT 1680
D 2646 TVTESDIRVEESIYCCDLAPEARQAIISLTERLYIGGLPLNSKQNGCYRRCRASGVLT 2705
QY 1681 TSCGNTLTLCYKAAACRAAKLQCTMLVCGDGLVWICESAGTORDEASLRFAETAMTRY 1740
D 2706 TSCGNTLTLCYKASACRAAKLQCTMLVCGDGLVWICESAGIQEDASLRVFTAMTRY 2765
QY 1741 SAPGDPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLAFAAWETARHTPVN 1800
D 2766 SAPGDPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLAFAAWETARHTPIN 2825
QY 1801 SWLGNIIIMYAPTLWARMTLMTHTFSSILLAQOLEKALDCQIYGACYSIEPLDLQIQL 1860
D 2826 SWLGNIIIMYAPTLWARMTLMTHTFSSILLAQOLEKALDCQIYGATYSIEPLDLQIQL 2885
QY 1861 HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWHRARSVRALLSQGGRATCGKYL 1920
D 2886 HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWHRARSVRALLSQGGRATCGKYL 2945
QY 1921 NNAVTKLKTPTPAASQOLDSSWFWAGYSGDDIYHLSLRARPRFWMCCLLLSVGVGIY 1980
D 2946 NNAVTKLKTPTPAASQOLDSSWFWAGYSGDDIYHLSLRARPRFWMCCLLLSVGVGIY 3005
QY 1981 LLENR 1985
D 3006 LLENR 3010

RESULT 13
P90193 PRELIMINARY; PRT: 3010 AA.
ID P90193
AC P90193;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HCV-1b;
RX MEDLINE=95340824; PubMed=7542279;
RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
RA Yamamoto C., Izumi N., Marumo F., Sato C.;
RT "Comparison of full-length sequences of interferon-sensitive and
RL resistant hepatitis C virus 1b.";
RN J. Clin. Invest. 96:224-230 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HCV-1b;
RA Enomoto N.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: D50485, BAA03076.1; -.
DR PIR: A61196; A61196.
DR PIR: PQ0246; PQ0246.
DR PIR: PS0329; PS0329.
DR HSSP: P26663; IUXP.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
```

```
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002531; HCV_env.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRP.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; Viral_RdRP; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT CHAIN 810 1026 NS2.
FT CHAIN 1027 1657 NS3.
FT CHAIN 1658 1711 NS4a.
FT CHAIN 1 191 core protein.
FT CHAIN 1712 1972 NS4b.
FT CHAIN 1973 2419 NS4b.
FT CHAIN 2420 3010 NSSB.
FT CHAIN 192 383 E1.
FT CHAIN 384 809 E2.
SQ SEQUENCE 3010 AA; 326886 MW; 21CD35B3DAC02B84 CRC64;

Query Match 97.3%; Score 10182; DB 2; Length 3010;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1918; Conservative 34; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCIITSLTGRBNQVEGVQVVSATQSFATCNGVCTVYHGA 60
D 1026 LAPITAYSQOTRGLGCIITSLTGRDNQVEGVQVVSATQSFATCNGVCTVYHGA 1085
QY 61 GSKTLGPKGPIQMTYNTVDQDLVGMQAPPGARSLTPCTCGSSDLVLTTHADVIPVRR 120
D 1086 GSKTLGSKGPIQMTYNTVDQDLVGMQAPPGARSLTPCTCGSSDLVLTTHADVIPVRR 1145
QY 121 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
D 1146 GDGRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVSMETTM 1205
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTSKSTKVPAAVAAQGYKVLNPNVAATLFGF 240
D 1206 RSEVFTDNSPPAVPQTFQVAHLHAPTSKSTKVPAAVAAQGYKVLNPNVAATLFGF 1265
QY 241 AYMSKAHGIDPNIRTCVTRITTGAPITYTYGKFLADGGCGSGAYDI1ICDCCHSTDST 300
D 1266 AYMSKAHGVDPNIRTCVTRITTGAPITYTYGKFLADGGCGSGAYDI1ICDCCHSTDST 1325
```

QY	301	ILIGTIVLDOAETAGARLWVLAATATPPGSVTPVPHENIEEVALSSTGCEIPFYKAIPIETI	360	QY	1381	STVSEASEDGVCCSMTWTGALITPCAAEETKLPINALNSLRLHNLVYATTSSAS	1440
DB	1326	ILIGTIVLDOAETAGARLWVLAATATPPGSVTPVPHENIEEVALSNTGEIPFYKAIPIETI	1385	DB	2406	STVSEASEDGVCCSMTWTGALITPCAAEESKLPINALNSLRLHNLVYATTSSAS	2465
QY	361	KGRHLIFCHSKKKKDELAALKSLGLNANAVYRGLDVSVIPTSQDVIWATDALMTGPT	420	QY	1441	LRQKVTTRDLQVLDHDDYRDLVKEMKAKASTVKALLSVEEACKLTTPPHSARSKFGYGA	1500
DB	1386	KGRHLIFCHSKKKKDELAALKSLGLNANAVYRGLDVSVIPTSQDVIWATDALMTGPT	1445	DB	2466	QRQKVTTRDLQVLDHDDYRDLVKEMKAKASTVKALLSVEEACKLTTPPHSARSKFGYGA	2525
QY	421	GFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGMRGIYRFTPG	480	QY	1501	DYRNLSKAVNHIRSVWKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVPPDLG	1560
DB	1446	GFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGMRGIYRFTPG	1505	DB	2526	DYRNLSKAVNHIRSVWKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVPPDLG	2585
QY	481	ERPSGMFSSVLCBCVDACANVELTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL	540	QY	1561	VAVCEKMALYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS	1620
DB	1506	ERPSGMFSSVLCBCVDACANVELTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL	1565	DB	2586	VAVCEKMALYDVVSTLPOAVMGPSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS	2645
QY	541	THIDAFHLSQTKOAGDNFFYLVAQATVCARAQAPPPSDMWKCLIRLKPILHGTPLL	600	QY	1621	TVTENDIRVEESIYOCDDLAPPEARQAIRSLTBRLYIGPLTNSKQNCYRRCRASGVLT	1680
DB	1566	THIDAFHLSQTKOAGDNFFYLVAQATVCARAQAPPPSDMWKCLIRLKPILHGTPLL	1625	DB	2646	TVTESDIRVEESIYOCDDLAPPEARQAIRSLTBRLYIGPLTNSKQNCYRRCRASGVLT	2705
QY	601	YRLGAVQNEVTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR	660	QY	1681	TSCGNLTLCYLAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY	1740
DB	1626	YRLGAVQNEVTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR	1685	DB	2706	TSCGNLTLCYLAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY	2765
QY	661	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKALGLOTATKQAEA	720	QY	1741	SAPPDGPPEYDLELITSCSSNSVVAHDASGRVYVYLTTRDPTTPLARAWEATARTPVN	1800
DB	1686	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKALGLOTATKQAEA	1745	DB	2766	SAPPDGPPEYDLELITSCSSNSVVAHDASGRVYVYLTTRDPTTPLARAWEATARTPVN	2825
QY	721	AAPVVEKWRTEAFWAKHWNFIISGIQVLAGLSTLPGNPAIASLMAFTASITSPITTOH	780	QY	1801	SWLGNIMVAPTILWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPQIIQRL	1860
DB	1746	AAPVVEKWRTEAFWAKHWNFIISGIQVLAGLSTLPGNPAIASLMAFTASITSPITTOH	1805	DB	2826	SWLGNIMVAPTILWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPQIIQRL	2885
QY	781	TLLFNILGCVAAQALAPPASAFVAGIAGAAVSGIGLKVLDVILAGYAGAGALVA	840	QY	1861	HGLSFAFSLSHSPGEINRVASCLRKLGVPLRVWRHRSVRAARLSQGGRAATCGKLYF	1920
DB	1806	TLLFNILGCVAAQALAPPASAFVAGIAGAAVSGIGLKVLDVILAGYAGAGALVA	1865	DB	2886	HGLSFAFSLSHSPGEINRVASCLRKLGVPLRVWRHRSVRAARLSQGGRAATCGKLYF	2945
QY	841	FKVMSGEMSTEDLVNLLPAILSPGALVGVVCAAILRHHVGPGEAGVOMNKLIAFASR	900	QY	1921	RNAVRTKLTPIPAASQDLSSWVAGYSGGDIYHLSRARPRWPMWMLLLSVGVGIY	1980
DB	1866	FKVMSGEMSTEDLVNLLPAILSPGALVGVVCAAILRHHVGPGEAGVOMNKLIAFASR	1925	DB	2946	RNAVRTKLTPIPAASQDLSSWVAGYSGGDIYHLSRARPRWPMWMLLLSVGVGIY	3005
QY	901	GNHVSPTHVPSDAAARVQILSSLTITQLKRLHQWINECDSTPCSGSLRDVMDWIC	960	QY	1981	LLPNR 1985	
DB	1926	GNHVSPTHVPSDAAARVQILSSLTITQLKRLHQWINECDSTPCSGSLRDVMDWIC	1985	DB	3006	LLPNR 3010	
QY	961	TVLTDFTWMLQSKLLPRLPGVPFFSCQRYKGVWVGDMQITTCPCQAQITGHVKGSMR	1020	RESULT 14			
DB	1986	TVLTDFTWMLQSKLLPRLPGVPFFSCQRYKGVWVGDMQITTCPCQAQITGHVKGSMR	2045	Q9J3H3	PRELIMINARY;	PRT; 3010 AA.	
QY	1021	IYGPRTCSNTHGCTEPINAVTTGCTPSPAPNYSRALRVAAEYVEVTRVGDVHYVTGM	1080	AC	Q9J3H3;		
DB	2046	IYGPRTCSNTHGCTEPINAVTTGCTPSPAPNYSRALRVAAEYVEVTRVGDVHYVTGM	2105	DT	01-OCT-2000 (T-EMBLrel. 15, Created)		
QY	1081	TTDNVPCQVPAPEFFTEVDGVRHLRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE	1140	DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)		
DB	2106	TTDNVPCQVPAPEFFTEVDGVRHLRYAPACKPLLEEVTVFLVGLNQYLVGSQLPCEPE	2165	DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)		
QY	1141	PDVAVITSMITDPSHITAEAKRLARGSPPSIASSASQLSAPSLKATCTTHDSDPAD	1200	DE	Polyprotein.		
DB	2166	PDVAVITSMITDPSHITAEAKRLARGSPPSIASSASQLSAPSLKATCTTHDSDPAD	2225	OS	Hepatitis C virus.		
QY	1201	LIEANILLWQEMGNITRVSEKNVWILDSFEPLQAEDEEREVSVAEILRRSRKTPPRAM	1260	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
DB	2226	LIEANILLWQEMGNITRVSEKNVWILDSFEPLQAEDEEREVSVAEILRRSRKTPPRAM	2285	OC	Hepacivirus.		
QY	1261	PIWAPDYNPLLESWKDPDYPVPHVHCPLPAPAKPIPPPRKRTVWLSSTVSALA	1320	OX	[1]_TaxID=11103;		
DB	2286	PIWAPDYNPLLESWKDPDYPVPHVHCPLPAPAKPIPPPRKRTVWLSSTVSALA	2345	RN	SEQUENCE FROM N.A.		
QY	1321	ELATKTFGSSSASVDSGTATASDPQSDDGAGSDVESYSSPPLLEGEEDPLSDGSW	1380	RC	STRAIN=MD19;		
DB	2346	ELATKTFGSSSASVDSGTATASDPQSDDGAGSDVESYSSPPLLEGEEDPLSDGSW	2405	RA	Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;		

DR GO: 0008236; F: serine-type peptidase activity; IEA.  
 DR GO: 0005198; F: structural molecule activity; IEA.  
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.  
 DR GO: 0006350; P: transcription; IEA.  
 DR GO: 0019079; P: viral genome replication; IEA.  
 DR GO: 0019087; P: viral transformation; IEA.  
 DR InterPro: IPR000345; CytC heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR002522; HCV capsid.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRP.  
 DR InterPro: IPR001650; Helicase C.  
 DR InterPro: IPR004109; Peptidase S29.  
 DR InterPro: IPR009003; Pept\_Ser\_Cys.  
 DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR007094; RNA\_pol\_PSVir.  
 DR Pfam: PF01543; HCV capsid; 1.  
 DR Pfam: PF01542; HCV core; 1.  
 DR Pfam: PF01539; HCV env; 1.  
 DR Pfam: PF01560; HCV NS1; 1.  
 DR Pfam: PF01538; HCV NS3; 1.  
 DR Pfam: PF02907; HCV NS2; 1.  
 DR Pfam: PF01006; HCV NS4a; 1.  
 DR Pfam: PF01001; HCV NS4b; 1.  
 DR Pfam: PF01506; HCV NS5a; 1.  
 DR Pfam: PF00271; Helicase C; 1.  
 DR Pfam: PF00998; Viral RdRP; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 SQ SEQUENCE 3010 AA; 327235 MW; 44C34677649CB8DD CRC64;

Query Match 97.2%; Score 10177; DB 2; Length 3010;  
 Best Local Similarity 96.1%; Pred. No. 0;  
 Matches 1908; Conservative 44; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLLCITSLTGRDNQVEGVQVSTATQSFATCNGVCWTVYHGA 60  
 DB 1026 LAPITAYSQOTRGLLCITSLTGRDNQVEGVQVSTATQSFATCNGVCWTVYHGA 1085

QY 61 GSXKLAGPKGPIQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLXLVTRHADVIPVRR 120  
 DB 1086 GAKTLAGPKGPIQMYTNVDQDLVGWQSPGARSLETCGSSDLXLVTRHADVIPVRR 1145

QY 121 GDSRGLSPRPVSYLKGSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
 DB 1146 GDSRGLSPRPISYLYLKGSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSEMETTM 1205

QY 181 RSPVFTDNSSPRAVPTQFVAHLHAPTGSCKSTKVAAYAAQYKVLNPSVAATLGF 240  
 DB 1206 RGFVFTDNSSPRAVPTQFVAHLHAPTGSCKSTKVAAYAAQYKVLNPSVAATLGF 1265

QY 241 AYMSKAHGIDPNIRTVRTITGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
 DB 1266 AYMSKAHGIEPSIRTVRTITGAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 1325

QY 301 ILGIGTVLQOATAGARLVVLTATPPGVSVPVPHNIEEVALSSGTGEIPFYKAIPIETI 360  
 DB 1326 ILGIGTVLQOATAGARLVVLTATPPGVSVPVPHNIEEVALSSGTGEIPFYKAIPIEI 1385

QY 361 KGRGHLIFCHSKKKCKDELAAGSLGNAVAYVYRGLDVSVIPTSGDIVVATDALTMTGT 420  
 DB 1386 KGRGHLIFCHSKKKCKDELAAGSLGNAVAYVYRGLDVSVIPASGGVVVATDALTMTGT 1445

QY 421 GQFDSVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRQRGRGTGRGMYRFTVPG 480  
 DB 2526 DVRSLSRAINHIHWSVKDLLEDTEPTTQTTIMAKNEVFCVQPEKGRKPARLI VFPDGL 2585

DB 1446 GQFDSVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRQRGRGTGRGMYRFTVPG 1505  
 QY 481 ERPSGMFDDSSVLCBCEYDAGCAWYELTAPETSRLRAYINTPLPVCODHLEFWESVETGL 540  
 DB 1506 ERPSGMFDDSSVLCBCEYDAGCAWYELTAPETSRLRAYINTPLPVCODHLEFWESVETGL 1555

QY 541 THIDAHFLSQTQKAGDNFPYLVAQATVCARAQAPPPSWDMKCLIRLKP TLHGPTPL 600  
 DB 1566 THIDAHFLSQTQKAGDNFPYLVAQATVCARAQAPPPSWDMKCLIRLKP TLHGPTPL 1625

QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTWLVGVGLAALAAYCLTGTGSVWLVGR 660  
 DB 1626 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTWLVGVGLAALAAYCLTGTGSVWLVGR 1685

QY 661 IILSGKPAIIPDREVLYREFDEMECASHLPYIEQGMQLAEQFKQKAIQLQTATQAEA 720  
 DB 1686 IILSGKPAIIPDREVLYREFDEMECASHLPYIEQGMQLAEQFKQKAIQLQTATQAEA 1745

QY 721 AAPVSESKRTLEAFWAKHWNFI SGIQYIAGLSTLPGNPAIASLMFTAFTSITPLTQH 780  
 DB 1746 AAPVSESKRTLEAFWAKHWNFI SGIQYIAGLSTLPGNPAIASLMFTAFTSITPLTQH 1805

QY 781 TLLFNILGGWVAQAAPPASAASAFVAGIAGAVSGTGLGKVLVDILAGYGAGVAGALVA 840  
 DB 1806 TLLFNILGGWVAQAAPPASAASAFVAGIAGAVSGTGLGKVLVDILAGYGAGVAGALVA 1865

QY 841 FKVMGEMFSTEDLVNLLPAILSPALVGVVCAAILRRHVGPGEVAVMMNLLIAFASR 900  
 DB 1866 FKVMGEMFSTEDLVNLLPAILSPALVGVVCAAILRRHVGPGEVAVMMNLLIAFASR 1925

QY 901 GNVHSTHYVPESDAARVTOILSSITITQLKRLHWNEDCSTPCSGSWLRDWDWIC 960  
 DB 1926 GNVHSTHYVPESDAARVTOILSSITITQLKRLHWNEDCSTPCSGSWLRDWDWIC 1985

QY 961 TVLTDFTKLQSKLLPRLPGVPFPFCORGKGVWVGDMQTTCPGCAOITGHVKNKSMR 1020  
 DB 1986 TVLADFTKLQSKVLPRLPGVPFPFCORGKGVWVGDMQTTCPGCAHIAHVNKSMR 2045

QY 1021 IVGPRTCSNTWHGTFFINAYTTGPTCPSPAPNYSRALWRVAEEYVEVTRVGVPHVVTGM 1080  
 DB 2046 IVGPRTCSNTWHGTFFINAYTTGPTCPSPAPNYSRALWRVAEEYVEITFVGDHVTGM 2105

QY 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPLLRREVTFLVGNQYLVGSQLPCEPE 1140  
 DB 2106 TTDNIKCPQVPAPEFFTELDGVLHRYAPACKPLLRREVTFOVGLNQYAVGSQLPCEPE 2165

QY 1141 PDVAVLTSLMTDPSHITTAETAKRLARGSPPSLIASSASQLSAPSUKATCTTHDSDPAD 1200  
 DB 2166 PDVAVLTSLMTDPSHITTAETAKRLARGSPPSLIASSASQLSAPSUKATCTTHDSDPAD 2225

QY 1201 LIEANLLWQEMGNNITRVESENKVVILDSFFLQAEDEEREVSVPAEILRKSKEPPRAM 1260  
 DB 2226 LIEANLLWQEMGNNITRVESENKVVILDSFFLQAEDEEREVSVPAEILRKSKEPPRAM 2285

QY 1261 PIWARPDPNPPLLESKWDPDYPPVHGCPLPAPAKAPPIPPPRKRRTVLSSTVSSALA 1320  
 DB 2286 PIWARPDPNPPLLESKWDPDYPPVHGCPLPAPAKAPPIPPPRKRRTVLSSTVSSALA 2345

QY 1321 ELATKTFGSSSADVSDGTATASPOQPSDDGAGSDVESYSSMPPLEGGPDLSDGWS 1380  
 DB 2346 ELATKTFGSSSADVSDGTATASPOQPSDDGAGSDVESYSSMPPLEGGPDLSDGWS 2405

QY 1381 STVSEASEDDVCCSMSTYTWGALITPCAAEETKLPINALNSLLRHHNLVYATTSRSAS 1440  
 DB 2406 STVSEASEDDVCCSMSTYTWGALITPCAAEESKLPINALNSLLRHHNLVYATTSRSAS 2465

QY 1441 LRQKVTDFRLQVLDHHDVRLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 1500  
 DB 2466 QRQKVTFDLQVLDHHDVRLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 2525

QY 1501 DVNRLSSKAVNHIRSVKDLLEDTEPTTIDTTIMAKNEVFCVQPEKGRKPARLI VFPDGL 1560  
 DB 2526 DVRSLSRAINHIHWSVKDLLEDTEPTTIDTTIMAKNEVFCVQPEKGRKPARLI VFPDGL 2585



```
QY 1561 VRVCEKALYDVVSTLPQAVNGSSYGFQYSPQORVEFLVNAWAKKCPMGFYDTRCFDS 1620
Db 2586 VRVCEKALYDVVSTLPQAVNGASYGFQYSPQORVEFLVNAWKKCPMGFSYDTRCFDS 2645
QY 1621 TVTENDIIVERSIYOCCLAPEARQAIISLTERLYIGGLPLTNSKQNGYRRCRASGVLT 1680
Db 2646 TVTENDIITESIYOCCLAPARQAIISLTERLYIGGLPLNSKQNGYRRCRASGVLT 2705
QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740
Db 2706 TSCGNLTLCYLKATAACRAAKLQDCTMLVCGDDLVIICESAGTQDEAANLRFTEAMTRY 2765
QY 1741 SAPPGDPKPEYDLELITSCSSNVSAHDASGRVYVLTTRDPTTFLARAAMETARTHPVN 1800
Db 2766 SAPPGDPKPEYDLELITSCSSNVSAHDASGRVYVLTTRDPTTFLARAAMETARTHPVN 2825
QY 1801 SWLGNIMYAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 1860
Db 2826 SWLGNIMYAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 2885
QY 1861 HGLSAFSLHSYSPGBINRVASCLRLKGLVPPLEVRHRSVRARLLSQGRAATCGKYL 1920
Db 2886 HGLSAFSLHSYSPGBINRVASCLRLKGLVPPLEVRHRSVRARLLSQGRAATCGKYL 2945
QY 1921 NWAETKLTLPIDPAASQDLSSVFWAGYSGGDIYHLSRARPFRMFWCLLLLSVGVGII 1980
Db 2946 NWAETKLTLPIDPAAYQLDLSGWFVAGYSGGDIYHSVRARPFRMFWCLLLLSVGVGII 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 15
Q9QIX6 PRELIMINARY; PRT; 3010 AA.
AC Q9QIX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD8-1;
RX MEDLINE=20013325; PubMed=10544098;
RA Nagayama K., Kurotsaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
DR EMBL; AF165059; AAD56194.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; P26663; IJXP.
DR MEROPS; S29_002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
```

```
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol Psvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327468 MW; 4613744EC4D4A013 CRC64;

Query Match 97.2%; Score 10177; DB 2; Length 3010;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1916; Conservative 41; Mismatches 28; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCGIITSLTGRDRNQVEGEVQVSTATQSFATCNGVCWTYYHGA 60
Db 1026 LAPITAYSQOTRGLGCGIITSLTGRDRKNQVEGEVQVSTATQSFATCNGVCWTYYHGA 1085

QY 61 GSKTLAGPKGITQMYTNVDQDLVQWQAPPQPGARSITPTCTGSSDLYLVTRHADVIPVRR 120
Db 1086 GSKTLAGPKGITQMYTNVDQDLVQWQAPPQPGARSITPTCTGSSDLYLVTRHADVIPVRR 1145

QY 121 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db 1146 GDGRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVSMETTM 1205

QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTQSGSKTKVPAAYAAQGYKVLVLPNSVAATLGRG 240
Db 1206 RSPVFTDNSSPPAVPQTQVAHLHAPTQSGSKTKVPAAYAAQGYKVLVLPNSVAATLGRG 1265

QY 241 AYMSKAHGIDPNIRTVGRTITTTGAPITYSTYKGLADGCGSGAGYDIIICDCHSDTST 300
Db 1266 AYMSKAYGTDPNIRTVGRTITTTGAPITYSTYKGLADGCGSGAGYDIIICDCHSDTST 1325

QY 301 ILIGIGTVLDQAETAGARLVVLATATPPGSSVTVPHEINIEVALSSTGEIPFYKALPIETI 360
Db 1326 ILIGIGTVLDQAETAGARLVVLATATPPGSSVTVPHEINIEVALSSTGEIPFYKALPIETI 1385

QY 361 KGRHLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIPTSQDVIVVATDALMTGT 420
Db 1386 KGRHLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIPTSQDVIVVATDALMTGT 1445

QY 421 GPDFSDVDCNTCTVQTVDFSLDPTFTTETTTVPQAVSRQRGRGTMGYRFTVTPG 480
Db 1446 GPDFSDVDCNTCTVQTVDFSLDPTFTTETTTVPQAVSRQRGRGTMGYRFTVTPG 1505

QY 481 ERPSGMFDSSVLCECYDAGCAWYELTPAETSRLRAYLNTPLGLPVCQDHLRFWESVFTGL 540
```

```
Db 1506 ERPSGNFSDSVLCECYDAGCANYELTPAETSVURLAYLNTPLVQCQDHLHFWESVFTGL 1565
QY 541 THIDAFLSQTQAGDNFPYLVAYQATVCARAQPPSWDMWKCLIRLKPTLHGPTPL 600
Db 1566 THIDAFLSQTQAGDNFPYLVAYQATVCARAQPPSWDMWKCLIRLKPTLHGPTPL 1625
QY 601 YRLGATVONEVTHPTITKYMACHMSADLEVVTSWLVGVGLAALAAAYCLTTGSSVVIVGR 660
Db 1626 YRLGATVONEVTHPTITKYMACHMSADLEVVTSWLVGVGLAALAAAYCLTTGSSVVIVGR 1685
QY 661 IILSGKPAIIPDREVLIRFDEMEECASHLPYIEQGMQLABQFKOKAIGLLOATKQAEA 720
Db 1686 IILSGKPAIIPDREVLIRFDEMEECASHLPYIEQGMQLABQFKOKAIGLLOATKQAEA 1745
QY 721 AAPVESKVRTUEAFWAKHWNFIISGQYLAGLSTLPGNPAIASLMAFTASITPTTQH 780
Db 1746 AAPVESKVRTUEAFWAKHWNFIISGQYLAGLSTLPGNPAIASLMAFTASITPTTQH 1805
QY 781 TLLENILGGWVAOALAPPSAASAFVAGTAGIAGVSGIGLKVLDVILAGYGAGVAGALVA 840
Db 1806 TLLENILGGWVAOALAPPSAASAFVAGTAGIAGVSGIGLKVLDVILAGYGAGVAGALVA 1865
QY 841 FKVMSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVOMNRLIAFASR 900
Db 1866 FKVMSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVOMNRLIAFASR 1925
QY 901 GNVHSTHYVPESDAAARVTOILSSITITQLKRLHQMINECDSTPCSGSWLRDWDWIC 960
Db 1926 GNVHSTHYVPESDAAARVTOILSSITITQLKRLHQMINECDSTPCSGSWLRDWDWIC 1985
QY 961 TVLTDFKTLQSKLRLPGVFPFSCQYKGVWRGDMQITPCGQAOITGHVKNWSR 1020
Db 1986 TVLTDFKTLQSKLRLPGVFPFSCQYKGVWRGDMQITPCGQAOITGHVKNWSR 2045
QY 1021 IVGPRTCSNTWGTGTFPINAYTTGCTPPAPNYSRALWVAABEYVEVTRVGDVDFHYVTGM 1080
Db 2046 IVGPRTCSNTWGTGTFPINAYTTGCTPPAPNYSRALWVAABEYVEVTRVGDVDFHYVTGM 2105
QY 1081 TTDNVKPCQVPAPPEFTVGVRLHRYAPACKPLREBVTFLVGLNQYLVGSQLPCEPE 1140
Db 2106 TTDNVKPCQVPAPPEFTVGVRLHRYAPACKPLREBVTFLVGLNQYLVGSQLPCEPE 2165
QY 1141 PDVAVLTSMLTDPHSHTATKRLARGSPPSIASSASQLSAPSJKATCTTHDSPDAD 1200
Db 2166 PDVAVLTSMLTDPHSHTATKRLARGSPPSIASSASQLSAPSJKATCTTHDSPDAD 2225
QY 1201 LIEANLLWQEMGNITRVESENKVVILDSFBPQAEDEDEREVSVAEILRRSRKFPFRAM 1260
Db 2226 LIEANLLWQEMGNITRVESENKVVILDSFBPQAEDEDEREVSVAEILRRSRKFPFRAM 2285
QY 1261 PIWARPDPNPPLESWKDPDYVPVHVHGCPLPPAKAPPIPPRRKRKTVVLSSTVSSALA 1320
Db 2286 PIWARPDPNPPLESWKDPDYVPVHVHGCPLPPAKAPPIPPRRKRKTVVLSSTVSSALA 2345
QY 1321 ELATKTFGSSSANDSGTATASDPQSDGAGSVESYSMPPLEGEPPDLSDGSW 1380
Db 2346 ELATKTFGSSSANDSGTATASDPQSDGAGSVESYSMPPLEGEPPDLSDGSW 2405
QY 1381 STVSEASDENVCCMSYTWGTALITPCAAEETKLPINALSNSLRHNLVYATTSSAS 1440
Db 2406 STVSEASDENVCCMSYTWGTALITPCAAEESKLPINPLSNSLRHNLVYATTSSAS 2465
QY 1441 LRQKKVTFORLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFCYGAK 1500
Db 2466 LRQKKVTFORLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFCYGAK 2525
QY 1501 DYNRLSSKAVNHIRSVMKOLLDETETPIDITIMAKNEVFCVQPEKGRKPARIIVFPDLG 1560
Db 2526 DYNRLSSKAVNHIRSVMKOLLDETETPIDITIMAKNEVFCVQPEKGRKPARIIVFPDLG 2585
QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQORVEFLYNWAKKCPMGFAVDTRCFDS 1620
```

Search completed: December 8, 2004, 12:32:06  
Job time : 281 secs

```
Db 2586 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQORVEFLYNWAKKCPMGFSYDTRCFDS 2645
QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRASGVL 1680
Db 2646 TVTESDIRTEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRASGVL 2705
QY 1681 TSCGNTLTCTYLKAAACRAAKLODCTMLVCGDDLVVICESAGTOEDEASIRAFTEAMTRY 1740
Db 2706 TSCGNTLTCTYLKASACRAAKLODCTMLVNGDDLVVICESAGTOEDAASIRVTEAMTRY 2765
QY 1741 SAPPDGPPEYDLELITSCSNVSVAHDAAGKRVVYLTRDPTTPLARAAMETARHTPVN 1800
Db 2766 SAPPDGPPEYDLELITSCSNVSVAHDAAGKRVVYLTRDPTTPLARAAMETARHTPVN 2825
QY 1801 SWLGNIIIMVAPTILWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIORL 1860
Db 2826 SWLGNIIIMVAPTILWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIORL 2885
QY 1861 HGLSAPLSHSYSPGEINRVASCLRKLGVPPPLVWRHARSVRARLLSCGRAATCGKYL 1920
Db 2886 HGLSAPLSHSYSPGEINRVASCLRKLGVPPPLVWRHARSVRARLLSCGRAATCGKYL 2945
QY 1921 NNAVRTKLKLTIPAAASQLDLSWFWAGYSGGDIYHSLSRARPRWPMWCLLILSVGVGIY 1980
Db 2946 NNAVRTKLKLTIPAAASQLDLSWFWAGYSGGDIYHSLSRARPRWPMWCLLILSVGVGIY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:18:55 ; Search time 32059 Seconds  
(without alignments)  
11784.441 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagccccgattgggggc.....ggcctcttcgagatcaagt 7989

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7987.4	100.0	7989	6	AR406043
2	7987.4	100.0	7989	6	AX036255 Sequence
3	7987.4	100.0	7989	12	SSE242652
4	7987.4	100.0	10690	6	AJ242652 Hepatitis
5	7965.4	99.7	8001	6	AX739972
6	7965.4	99.7	8001	6	AR406042
7	7965.4	99.7	8001	6	AX036254
8	7947.8	99.5	8001	12	SSE242654
9	7947.8	99.5	8001	6	AR406047
10	7946.2	99.5	8001	6	AX036259
11	7946.2	99.5	8001	6	AR406045
12	7939.8	99.4	8001	6	AX036257
13	7939.8	99.4	8001	6	AR406049
14	7610.4	95.3	12305	6	AX036261
15	7607.2	95.2	12315	6	AX937622
16	7329.4	91.7	8637	6	AX937621
17	7329.4	91.7	8637	6	AR406041
18	7329.4	91.7	8637	6	AX036253
19	7307.4	91.5	8649	6	AJ242651 Hepatitis
					AR406044

20	7307.4	91.5	8649	6	AX036256
21	7307.4	91.5	8649	12	SSE242653
22	7272.2	91.0	8639	6	AR488100
23	7272.2	91.0	8639	6	AX472291
24	7271.2	91.0	8638	6	AR488121
25	7271.2	91.0	8638	6	AX472314
26	7248.2	90.7	8638	6	AR488105
27	7248.2	90.7	8638	6	AX472297
28	7247.2	90.7	8638	6	AR488122
29	7247.2	90.7	8638	6	AX472315
30	7245.2	90.7	8642	6	AR488101
31	7245.2	90.7	8642	6	AX472292
32	7244	90.7	8638	6	AR488104
33	7244	90.7	8638	6	AX472296
34	7228.2	90.5	8648	6	AR488103
35	7228.2	90.5	8648	6	AX472295
36	7221	90.4	8643	6	AR488102
37	7221	90.4	8643	6	AX472294
38	6675.8	83.6	9658	14	AB119282
39	6185.2	77.4	9605	6	AX739971
40	6185.2	77.4	9605	14	HCJ238799
41	6185.2	77.4	11076	6	AR406040
42	6185.2	77.4	11076	6	AX036252
43	6167.6	77.2	11076	6	AR406048
44	6167.6	77.2	11076	6	AX036260
45	6166	77.2	11076	6	AR406046

## ALIGNMENTS

RESULT 1  
AR406043  
LOCUS AR406043 7989 bp DNA  
DEFINITION Sequence 10 from patent US 6630343.  
ACCESSION AR406043  
VERSION AR406043.1 GI:40155170  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 7989)  
AUTHORS Bartenschlager,R.  
TITLE Hepatitis C virus culture system  
JOURNAL Patent: US 6630343-A 10 07-OCT-2003;  
FEATURES Location/Qualifiers  
source 1..7989  
/organism="unknown"  
/mol\_type="genomic DNA"

linear PAT 18-DEC-2003

Query Match	100.0%;	Score	7987.4;	DB	6;	Length	7989;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	7989;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
Qy	1	GCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG	60				
Db	1	GCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG	60				
Qy	61	TCTTCACGCAAGACGCTCTAGCCATGGCGTTAGTATAGTGTGTCGACGCTCCAGGAC	120				
Db	61	TCTTCACGCAAGACGCTCTAGCCATGGCGTTAGTATAGTGTGTCGACGCTCCAGGAC	120				
Qy	121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGGTGAGTACACCGGAATTGCCAG	180				
Db	121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGGTGAGTACACCGGAATTGCCAG	180				
Qy	181	GACGACGGGTCTTTTCTTGGATCAACCCGCTCAATGCTCGGAGATTGGGCGTGCCCC	240				
Db	181	GACGACGGGTCTTTTCTTGGATCAACCCGCTCAATGCTCGGAGATTGGGCGTGCCCC	240				
Qy	241	GCAGACTGCTAGCCAGTAGTGTGGTTCGCGAAAGGCGCTTGTGTACTGCTGTATAGG	300				





Pb 4621 GAGGACTGCTCCACGCCATGCTCCGGCTCGTGCTTAAGAGATGTTTGGATTGGATATGC 4680  
 QY 4681 ACGGTGTTGACTGATTTCAAGACCTTGGCTCCAGTCCAGTCCCTGCGCGGATGTCGGGA 4740  
 Db 4681 ACGGTGTTGACTGATTTCAAGACCTTGGCTCCAGTCCAGTCCCTGCGCGGATGTCGGGA 4740  
 QY 4741 GTCCCTCTTCTTCATGTCACGTTGGTACAGGAGTCTGCGGGGCGACGGCATCATG 4800  
 Db 4741 GTCCCTCTTCTTCATGTCACGTTGGTACAGGAGTCTGCGGGGCGACGGCATCATG 4800  
 QY 4801 CAAACCACTGCTCCATGTCGAGCACAGATCACCGACATGTGAAACACGTTCCATGAGG 4860  
 Db 4801 CAAACCACTGCTCCATGTCGAGCACAGATCACCGACATGTGAAACACGTTCCATGAGG 4860  
 QY 4861 APTGTTGGGCTTACGACCTGTAGTAACACGTTGGCATGGAACATTTCCCAATTAAAGCGTAC 4920  
 Db 4861 APTGTTGGGCTTACGACCTGTAGTAACACGTTGGCATGGAACATTTCCCAATTAAAGCGTAC 4920  
 QY 4921 ACCAGGGCCCTGACGCGCTCCCGGGCCCAAAATTATCTAGGCGCTGTGGCGGGTG 4980  
 Db 4921 ACCAGGGCCCTGACGCGCTCCCGGGCCCAAAATTATCTAGGCGCTGTGGCGGGTG 4980  
 QY 4981 GTGCTGAGGAGTACGCTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGCGCATG 5040  
 Db 4981 GTGCTGAGGAGTACGCTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGCGCATG 5040  
 QY 5041 ACCACTGACAAAGTAAAGTCCCGTGTACAGTTCCCGGCCCGGATTTCTTACAGAAAGTG 5100  
 Db 5041 ACCACTGACAAAGTAAAGTCCCGTGTACAGTTCCCGGCCCGGATTTCTTACAGAAAGTG 5100  
 QY 5101 GATGGGTGGGTGACAGGATGCTCCAGGTGCAAAACCCCTCTACGGGAGGAGTTC 5160  
 Db 5101 GATGGGTGGGTGACAGGATGCTCCAGGTGCAAAACCCCTCTACGGGAGGAGTTC 5160  
 QY 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTGGGTACAGTCCCATGCGGACCGGAA 5220  
 Db 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTGGGTACAGTCCCATGCGGACCGGAA 5220  
 QY 5221 CCGGACGTAGAGTGTCACTTCATGCTCACGACCCCTCCACATTAACGGCGGAGACG 5280  
 Db 5221 CCGGACGTAGAGTGTCACTTCATGCTCACGACCCCTCCACATTAACGGCGGAGACG 5280  
 QY 5281 GCTAAGGTAGGTGCGCAGGGATCTCCCGCTCTTGGCTGCTCCAGTACAGTACGAG 5340  
 Db 5281 GCTAAGGTAGGTGCGCAGGGATCTCCCGCTCTTGGCTGCTCCAGTACAGTACGAG 5340  
 QY 5341 CTGCTGCGCTTCTTTGAAGGCAACATGCACTACCGTACATGCTACCGTACCGGACGCTGAC 5400  
 Db 5341 CTGCTGCGCTTCTTTGAAGGCAACATGCACTACCGTACATGCTACCGTACCGGACGCTGAC 5400  
 QY 5401 CTGCTGCGCTTCTTTGAAGGCAACATGCACTACCGTACATGCTACCGTACCGGACGCTGAC 5460  
 Db 5401 CTGCTGCGCTTCTTTGAAGGCAACATGCACTACCGTACATGCTACCGTACCGGACGCTGAC 5460  
 QY 5461 TCAGAAATAGGTAGTAAATTTGGACTTTTCGAGCGCTTCCAGCGGAGGAGATGAG 5520  
 Db 5461 TCAGAAATAGGTAGTAAATTTGGACTTTTCGAGCGCTTCCAGCGGAGGAGATGAG 5520  
 QY 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGGAAATTTCCCTCGAGCGATG 5580  
 Db 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGGAAATTTCCCTCGAGCGATG 5580  
 QY 5581 CCATATGGGACCGCGGATTAACACCTCCACTGTTAGTCTGAGGAGGAGGAGGAG 5640  
 Db 5581 CCATATGGGACCGCGGATTAACACCTCCACTGTTAGTCTGAGGAGGAGGAGGAGGAG 5640  
 QY 5641 TACGTCCTCCAGTGTACCGGTGTCATTGCGGCTGCAAGGCGCTCCGATACCA 5700  
 Db 5641 TACGTCCTCCAGTGTACCGGTGTCATTGCGGCTGCAAGGCGCTCCGATACCA 5700  
 QY 5701 CCTCCAGGAGGAGGAGCGGTTGCTGTCAGAAATCTACCGGTCTTCTGCTGCGG 5760  
 Db 5701 CCTCCAGGAGGAGGAGCGGTTGCTGTCAGAAATCTACCGGTCTTCTGCTGCGG 5760

QY 5761 GAGCTCGCCACAAAGACCTTTGGGAGCTCCGAAATCGTGGCGCTCGACAGCGGCA 5820  
 Db 5761 GAGCTCGCCACAAAGACCTTTGGGAGCTCCGAAATCGTGGCGCTCGACAGCGGCA 5820  
 QY 5821 ACGGCTCTCTCTGACAGCCCTCCAGACGCGGAGTCCGACGTTGAGTCTGATC 5880  
 Db 5821 ACGGCTCTCTCTGACAGCCCTCCAGACGCGGAGTCCGACGTTGAGTCTGATC 5880  
 QY 5881 TCCTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCGAGGCTTGG 5940  
 Db 5881 TCCTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCGAGGCTTGG 5940  
 QY 5941 TCCTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCGAGGCTTGG 6000  
 Db 5941 TCCTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCGAGGCTTGG 6000  
 QY 6001 ACAGGCGCTTGTATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCTCAATGACTG 6060  
 Db 6001 ACAGGCGCTTGTATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCTCAATGACTG 6060  
 QY 6061 AGCAACTCTTGTCTCGGTACCAACACTTGGTCTATGCTACACATCTCGAGCGCAAGC 6120  
 Db 6061 AGCAACTCTTGTCTCGGTACCAACACTTGGTCTATGCTACACATCTCGAGCGCAAGC 6120  
 QY 6121 CTGCGGACAGAAAGGTACCTTTGACAGACTGCAAGTCTCTGAGACCACTACCGGAC 6180  
 Db 6121 CTGCGGACAGAAAGGTACCTTTGACAGACTGCAAGTCTCTGAGACCACTACCGGAC 6180  
 QY 6181 GTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGCTAAATCTTATCCGTGGAG 6240  
 Db 6181 GTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGCTAAATCTTATCCGTGGAG 6240  
 QY 6241 GAAGCCTGTAGTACGCGCCCAACATTCGCGCAGATCTAAATTTGCGTATGGGCAAG 6300  
 Db 6241 GAAGCCTGTAGTACGCGCCCAACATTCGCGCAGATCTAAATTTGCGTATGGGCAAG 6300  
 QY 6301 GAGCTCCGAAACCTATCAGCAAGGCGGTTAAACACATCCGCTCCGTGTGGAAGGACTTG 6360  
 Db 6301 GAGCTCCGAAACCTATCAGCAAGGCGGTTAAACACATCCGCTCCGTGTGGAAGGACTTG 6360  
 QY 6361 CTGGAAGACACTGAGACACCAATTTGACACCAATCATGCGCAAAATAGGTTTCTGC 6420  
 Db 6361 CTGGAAGACACTGAGACACCAATTTGACACCAATCATGCGCAAAATAGGTTTCTGC 6420  
 QY 6421 GTCCAAACAGAGAGGCGGCGCAGCAGCTGCGCTTATCGTATTTCCAGATTTGGGG 6480  
 Db 6421 GTCCAAACAGAGAGGCGGCGCAGCAGCTGCGCTTATCGTATTTCCAGATTTGGGG 6480  
 QY 6481 GTTCGTGTGTCGAGAAATGCGCCCTTTACGATGTGCTCTCCACCTCCCTCAGGCGGTG 6540  
 Db 6481 GTTCGTGTGTCGAGAAATGCGCCCTTTACGATGTGCTCTCCACCTCCCTCAGGCGGTG 6540  
 QY 6541 ATGCGCTCTTATACCGAATTCGAATCTCTCTGAGAGCGGCTCGAGTTCTGTTGAAT 6600  
 Db 6541 ATGCGCTCTTATACCGAATTCGAATCTCTCTGAGAGCGGCTCGAGTTCTGTTGAAT 6600  
 QY 6601 GCCTGGAACGAGAAATGCGCTTATGGCTTCGATATGACACCGCTCTTTTACTCA 6660  
 Db 6601 GCCTGGAACGAGAAATGCGCTTATGGCTTCGATATGACACCGCTCTTTTACTCA 6660  
 QY 6661 ACCGTCACTGAGAAATGACATCCGTTGTGAGGAGTCAATCTACCAATTTGTGACTGGCC 6720  
 Db 6661 ACCGTCACTGAGAAATGACATCCGTTGTGAGGAGTCAATCTACCAATTTGTGACTGGCC 6720  
 QY 6721 CCGGAGGACGAGAGGCGGATGAGTCTGCTCAGAGGCGGCTTTTACATCGGGGCGCCCTG 6780  
 Db 6721 CCGGAGGACGAGAGGCGGATGAGTCTGCTCAGAGGCGGCTTTTACATCGGGGCGCCCTG 6780  
 QY 6781 ACTAATTTCTAAAGGAGAACTGCGGCTATCGCGGTATCGCGGCTGCGGCGGAGGCTGACTGACG 6840  
 Db 6781 ACTAATTTCTAAAGGAGAACTGCGGCTATCGCGGTATCGCGGCTGCGGCGGAGGCTGACTGACG 6840



6841 ACCAGCTGGGTAATACCTCACAATGTTACTTTGAAGCGCGCTGCGGCGCTGTCGAGCTGCG 6900  
6841 ACCAGCTGGGTAATACCTCACAATGTTACTTTGAAGCGCGCTGCGGCGCTGTCGAGCTGCG 6900  
6901 AAGCTCCAGGACTGCAAGATCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAGC 6960  
6901 AAGCTCCAGGACTGCAAGATCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAGC 6960  
6961 GCGGGACCCAAAGAGGAGGAGCGAGCTACGGGCTTCAGGAGGCTATGACTAGATAC 7020  
6961 GCGGGACCCAAAGAGGAGGAGCGAGCTACGGGCTTCAGGAGGCTATGACTAGATAC 7020  
7021 TCTCCCGCCCTGCGGAGCCCGCCCAACAGAAATACGACTTGGAGTTGATAAATCATGC 7080  
7021 TCTCCCGCCCTGCGGAGCCCGCCCAACAGAAATACGACTTGGAGTTGATAAATCATGC 7080  
7081 TCTTCCAATGTGTCAGTGGCGACGATGCAATCTGGCAAAAGGGTGTACTATCTCACCCGT 7140  
7081 TCTTCCAATGTGTCAGTGGCGACGATGCAATCTGGCAAAAGGGTGTACTATCTCACCCGT 7140  
7141 GACCCACCAACCCCTTGGCGGGCTGCGTGGGAGACGATAGACACACTCCAGTCAAT 7200  
7141 GACCCACCAACCCCTTGGCGGGCTGCGTGGGAGACGATAGACACACTCCAGTCAAT 7200  
7201 TCTTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGGCAAGGATGATCCTGATG 7260  
7201 TCTTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGGCAAGGATGATCCTGATG 7260  
7261 ACTCAATTTCTTCCATCTTCTAGCTCAGGAAACACTTGAAGAGCCCTAGATTGTGAC 7320  
7261 ACTCAATTTCTTCCATCTTCTAGCTCAGGAAACACTTGAAGAGCCCTAGATTGTGAC 7320  
7321 ATCTACGGGCGCTTACTCCATGAGCAGCTGACCTACTGATGATCAATCAACGATC 7380  
7321 ATCTACGGGCGCTTACTCCATGAGCAGCTGACCTACTGATGATCAATCAACGATC 7380  
7381 CATGGCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
7381 CATGGCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
7441 TCAATGCTCAGGAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500  
7441 TCAATGCTCAGGAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500  
7501 GTCGCGCTAGGCTACTGTCAGGGGGAGGGCTGCCACTTGTGCAAGTACTCTTTC 7560  
7501 GTCGCGCTAGGCTACTGTCAGGGGGAGGGCTGCCACTTGTGCAAGTACTCTTTC 7560  
7561 AACTGGGCGAGTAAAGGACCAAGCTCAAACTCACTCCCAATCCCGGCTGCGTCCCAAGTTGGAT 7620  
7561 AACTGGGCGAGTAAAGGACCAAGCTCAAACTCACTCCCAATCCCGGCTGCGTCCCAAGTTGGAT 7620  
7621 TTAATCAGCTGGTTGTTGCTGTTTACAGCGGGGAGACATATATACAGCCCTGTCTCGT 7680  
7621 TTAATCAGCTGGTTGTTGCTGTTTACAGCGGGGAGACATATATACAGCCCTGTCTCGT 7680  
7681 GCCCGACCCCGCTGTTTCACTGCTGCTACTCTTCTTCTAGGGGTAGGATCTAT 7740  
7681 GCCCGACCCCGCTGTTTCACTGCTGCTACTCTTCTTCTAGGGGTAGGATCTAT 7740  
7741 CTACTCCCCAACGATGAAGCGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
7741 CTACTCCCCAACGATGAAGCGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
7861 TTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
7861 TTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
7921 TAGCTGTGAAGAGTCCGTGAGCGCTTGTACTGAGAGAGTGTGATCTAGGCTCTCTGC 7980

Db 7921 TAGCTGTGAAGAGTCCGTGAGCGCTTGAAGAGAGTGTGATGACTGCGCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989  
RESULT 2  
AX036255  
LOCUS AX036255 7989 bp DNA linear PAT 16-NOV-2000  
DEFINITION Sequence 4 from Patent EP1043399.  
ACCESSION AX036255  
VERSION AX036255.1 GI:11225871  
KEYWORDS Hepatitis C virus  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
REFERENCE 1  
AUTHORS Bartenschlager,R.D.  
TITLE Hepatitis c virus cell culture system  
JOURNAL Patent: Ep 1043399-A 4 11-OCT-2000;  
BARTENSCHLAGER RALF DR (DE)  
FEATURES  
source location/Qualifiers  
1. 7989  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
ORIGIN  
Query Match 100.0%; Score 7987.4; DB 6; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCCAGCCCCCGATTGGGGCGGACACTCCACATAGATCACTCCCTGTGAGGAACACTACTG 60  
Db 1 GCCAGCCCCCGATTGGGGCGGACACTCCACATAGATCACTCCCTGTGAGGAACACTACTG 60  
QY 61 TCTTACGCGAAGAGGCTCTAGCCATGGGCTTAGTATAGTGTCTGCGAACCGGTGAGTACACCGGAATTTGCCAG 120  
Db 61 TCTTACGCGAAGAGGCTCTAGCCATGGGCTTAGTATAGTGTCTGCGAACCGGTGAGTACACCGGAATTTGCCAG 120  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTTGCCAG 180  
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTTGCCAG 180  
QY 181 GAGACCGGGTCTCTTCTTGGATCAACCCGCTCAATGCCCTGGAGATTGGCGGTGCCCCC 240  
Db 181 GAGACCGGGTCTCTTCTTGGATCAACCCGCTCAATGCCCTGGAGATTGGCGGTGCCCCC 240  
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGTACTGCTGATAGG 300  
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGTACTGCTGATAGG 300  
QY 301 GTGCTTGGAGTGCCCGGAGGCTCTCGTAGACCGGTGACCAATGAGCAAGATCTCTAAAC 360  
Db 301 GTGCTTGGAGTGCCCGGAGGCTCTCGTAGACCGGTGACCAATGAGCAAGATCTCTAAAC 360  
QY 361 CTCAAAGAAACCAAGAGCGGCCCATGATTGAACAAGATGATGACGAGGTTCTC 420  
Db 361 CTCAAAGAAACCAAGAGCGGCCCATGATTGAACAAGATGATGACGAGGTTCTC 420  
QY 421 CGGCGCTTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGATGATGAGGCTGCT 480  
Db 421 CGGCGCTTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGATGATGAGGCTGCT 480  
QY 481 CTGATGCCCGCTGTTTCCGGCTGTGACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCCCGCTGTTTCCGGCTGTGACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCCGGTCCCTGAATGAATGACGAGGAGGCGGCTATCGTGGCTGGCCA 600

Db	541	ACCTGTCGGTGCCTGAACTGACGAGGAGGAGCGGGCTATCTGTCGGTGGCA	600	Qy	1681	ATGGGATCTGATCTGGGCTCGTGACATGCTTTACATGTTAGTCGAGGTTAAAA	1740
Qy	601	CGACGGGCTTCCCTTTCGCGAGCTGTCTCGAGTGTCTACTGAAGCGGAGGACTGCG	660	Db	1681	ATGGGATCTGATCTGGGGCTCGTGACATGCTTTACATGTTAGTCGAGGTTAAAA	1740
Db	601	CGACGGGCTTCCCTTTCGCGAGCTGTCTCGAGTGTCTACTGAAGCGGAGGACTGCG	660	Qy	1741	AACGTCCTAGGCCCCCGAAACACACGCGGAGCTGGTTTCTTTGAAAAAACACGATAATACC	1800
Qy	661	TGCTATTTGGGCGAAGTGGCGGGGAGGATCTCTGTATCTACCTTGTCTCTGCGGAGA	720	Db	1741	AACGTCCTAGGCCCCCGAAACACACGCGGAGCTGGTTTCTTTGAAAAAACACGATAATACC	1800
Db	661	TGCTATTTGGGCGAAGTGGCGGGGAGGATCTCTGTATCTACCTTGTCTCTGCGGAGA	720	Qy	1801	ATGGCGCTATTACGGCCTACTCCCAACAGACCGGAGGCTACTTTGGGTGCTCATCACT	1860
Qy	721	AAGTATCCATCATGCTGATCAATGCGCGGCTGCAATGCTGTGATCCGGTACTCTGCC	780	Db	1801	ATGGCGCTATTACGGCCTACTCCCAACAGACCGGAGGCTACTTTGGGTGCTCATCACT	1860
Db	721	AAGTATCCATCATGCTGATCAATGCGCGGCTGCAATGCTGTGATCCGGTACTCTGCC	780	Qy	1861	AGCTCACAGGCGGAGCAGGACAGCTCGAGGGGAGGTCCTAAGTGTCTCCACCGCA	1920
Qy	781	CATTTCGACCAACGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC	840	Db	1861	AGCTCACAGGCGGAGCAGGACAGCTCGAGGGGAGGTCCTAAGTGTCTCCACCGCA	1920
Db	781	CATTTCGACCAACGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC	840	Qy	1921	ACACAACTCTTCTCGCGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCC	1980
Qy	841	TTGTGATCAGGATGATCTGAGGAGAGCATCAGGGCTCGGCGAGCGAACTGTTCCG	900	Db	1921	ACACAACTCTTCTCGCGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCC	1980
Db	841	TTGTGATCAGGATGATCTGAGGAGAGCATCAGGGCTCGGCGAGCGAACTGTTCCG	900	Qy	1981	GGCTCAAAAGACCTTTGCGGCCCCAAAGGGCCCCAATCACCCAAATGTACCAATGTGGAC	2040
Qy	901	CCAGGCTCAAGCGCGCATGCGCGAGGAGATCTCGTGTGACCCCATGGCGATGCCCT	960	Db	1981	GGCTCAAAAGACCTTTGCGGCCCCAAAGGGCCCCAATCACCCAAATGTACCAATGTGGAC	2040
Db	901	CCAGGCTCAAGCGCGCATGCGCGAGGAGATCTCGTGTGACCCCATGGCGATGCCCT	960	Qy	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGGCGGTTCTTGACACATGCACCTGCG	2100
Qy	961	GCTTCCGGAATATCATGTTGGAATGCGCGCTTTCTCGGATTCATCGACTGTGGCGCG	1020	Db	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGGCGGTTCTTGACACATGCACCTGCG	2100
Db	961	GCTTCCGGAATATCATGTTGGAATGCGCGCTTTCTCGGATTCATCGACTGTGGCGCG	1020	Qy	2101	GGCAGCTCGGACCTTTACTTGTGTACGAGGATGCGATGTCAATTCGGGTGCGCGCGCG	2160
Qy	1021	TGGGTGTGGCGACCGCTATCAGGATAGCTGTGCTACCGTGTGATTTGCTGAGAGC	1080	Db	2101	GGCAGCTCGGACCTTTACTTGTGTACGAGGATGCGATGTCAATTCGGGTGCGCGCGCG	2160
Db	1021	TGGGTGTGGCGACCGCTATCAGGATAGCTGTGCTACCGTGTGATTTGCTGAGAGC	1080	Qy	2161	GGCGACAGAGGGGAGCCTACTCTCCCGCAGCGCGCTCTCTACTTTGAAGGCTCTTGG	2220
Qy	1081	TTGGCGGCAATGGGCTGACCGCTTCTCGTGTGCTTACCGTATCGCGCTCCCGATTGCG	1140	Db	2161	GGCGACAGAGGGGAGCCTACTCTCCCGCAGCGCGCTCTCTACTTTGAAGGCTCTTGG	2220
Db	1081	TTGGCGGCAATGGGCTGACCGCTTCTCGTGTGCTTACCGTATCGCGCTCCCGATTGCG	1140	Qy	2221	GGCGGTCACCTGCTGCCCCCTCGGGGACGCTGTGGGCTCTTTCCGGCTGCGCGGTGCG	2280
Qy	1141	AGCGCATCGCTTCTATCGCTTCTTTGACGAGTTCCTTGAGTTTAAACAGACCAACAG	1200	Db	2221	GGCGGTCACCTGCTGCCCCCTCGGGGACGCTGTGGGCTCTTTCCGGCTGCGCGGTGCG	2280
Db	1141	AGCGCATCGCTTCTATCGCTTCTTTGACGAGTTCCTTGAGTTTAAACAGACCAACAG	1200	Qy	2281	ACCCAGAGGGTTGCGAAGCGGTGGACTTTGTACCGCTCGAGTCTATGAAACCACTATG	2340
Qy	1201	GTTTCCCTTAGGGGATCAATTCGCGCCCTCCCGCTCCCGCTTACCGTTACTGGC	1260	Db	2281	ACCCAGAGGGTTGCGAAGCGGTGGACTTTGTACCGCTCGAGTCTATGAAACCACTATG	2340
Db	1201	GTTTCCCTTAGGGGATCAATTCGCGCCCTCCCGCTCCCGCTTACCGTTACTGGC	1260	Qy	2341	CGGTCCCGGCTTTCACGAGCAACTGCTCCCTCCCGCGGTACCCGAGACATTTCCAGGTG	2400
Qy	1261	CGAAGCGCTTGAATAAGCGCGGTGTGCTTGTCTATATGTTATTTCCACCATATTG	1320	Db	2341	CGGTCCCGGCTTTCACGAGCAACTGCTCCCTCCCGCGGTACCCGAGACATTTCCAGGTG	2400
Db	1261	CGAAGCGCTTGAATAAGCGCGGTGTGCTTGTCTATATGTTATTTCCACCATATTG	1320	Qy	2401	GCCCATCTACGCGCCCTACTGTAGCGCAAGAGCACTAAGGTGCGCGGTGGGTATGCA	2460
Qy	1321	CGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCGCTGTCTTTGAGCGAGCATTCCT	1380	Db	2401	GCCCATCTACGCGCCCTACTGTAGCGCAAGAGCACTAAGGTGCGCGGTGGGTATGCA	2460
Db	1321	CGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCGCTGTCTTTGAGCGAGCATTCCT	1380	Qy	2461	GCCCAAGGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCGCCGCCACCTTAGGTTTCGGG	2520
Qy	1381	AGGGGCTTTCCTCCCTCGCCAAAGGATGCAAGGTCTGTGAATGTGCTGAGGAGCA	1440	Db	2461	GCCCAAGGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCGCCGCCACCTTAGGTTTCGGG	2520
Db	1381	AGGGGCTTTCCTCCCTCGCCAAAGGATGCAAGGTCTGTGAATGTGCTGAGGAGCA	1440	Qy	2521	GGGTATATGTCTAAGGCACATGCTATCGACCCCTAACATCAGAACCCGGGGTAAAGCAATC	2580
Qy	1441	GTTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGACGCGAGCGG	1500	Db	2521	GGGTATATGTCTAAGGCACATGCTATCGACCCCTAACATCAGAACCCGGGGTAAAGCAATC	2580
Db	1441	GTTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGACGCGAGCGG	1500	Qy	2581	ACCAAGGGTGGCGGCTACGACTCTCACTATGGCAAGTTCCTTTCGCGACGGTGTGTC	2640
Qy	1501	AACCCCGCCTGGGAGAGGTGCTCTGGGCGCAAAAGCCAGCTGTATTAAGATACACCT	1560	Db	2581	ACCAAGGGTGGCGGCTACGACTCTCACTATGGCAAGTTCCTTTCGCGACGGTGTGTC	2640
Db	1501	AACCCCGCCTGGGAGAGGTGCTCTGGGCGCAAAAGCCAGCTGTATTAAGATACACCT	1560	Qy	2641	TCCTGGGGCGCTTATGACATCATATATGTATGATGAGTGGCCACTCAACTGACTCGACCACT	2700
Qy	1561	GCAAGGCGCAACACCCAGTGCACGTTGTGAGTTGATGTGGAAGAGTCAAA	1620	Db	2641	TCCTGGGGCGCTTATGACATCATATATGTATGATGAGTGGCCACTCAACTGACTCGACCACT	2700
Db	1561	GCAAGGCGCAACACCCAGTGCACGTTGTGAGTTGATGTGGAAGAGTCAAA	1620	Qy	2701	ATCTTGGGCATCGGACAGTCTCTGGCAAGCGGAGCGCTCGAGCGGCTCGCTGCTG	2760
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGGATCCCGAGAGGTACCCCATTTG	1680	Db	2701	ATCTTGGGCATCGGACAGTCTCTGGCAAGCGGAGCGCTCGAGCGGCTCGCTGCTG	2760
Db	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGGATCCCGAGAGGTACCCCATTTG	1680				

QY	2761	CTGCCACCGCTACGCTCGGGATCGGTCAACCGTCCCATCAATCAAAATCGAGAGGTG	2820
Db	2761		2820
QY	2821	GCTCTGTTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCATC	2880
Db	2821		2880
QY	2881	AAGGGGGAGGCACTCATTTCTGCCATTCCTCAAGAAAGAAATGTGATGAGCTCGCGCG	2940
Db	2881		2940
QY	2941	AAGCTGTCGGCTCGGACTCAATGTGTAGCATATTTACCGGGCTTGTATGATCCGTC	3000
Db	2941		3000
QY	3001	ATACCAACTAGCGAGACGTCATTGTGTAGCAACGAGCGCTCTAATGACGGGCTTTTACC	3060
Db	3001		3060
QY	3061	GGCGATTTGACTCGACTGCACTGCAATACATATGTGTCAACGAGCGCTCTAATGACGGGCTTTACC	3120
Db	3061		3120
QY	3121	CTGGACCGGCTTCACTTTCAGATTCAGACGACACCGTGCACAGACGGGTGTCACGCTCG	3180
Db	3121		3180
QY	3181	CAGCGCGAGGAGGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA	3240
Db	3181		3240
QY	3241	GAACGGCTTCGGGCAATGTTTCGATTCCTCGGTTCTGTGCGAGTCTATGACGGGGCTGT	3300
Db	3241		3300
QY	3301	GCTTGTGACAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA	3360
Db	3301		3360
QY	3361	CCAGGTTGCCCTCTGCCAGGACCACTTCGAGTTCTGGGAGAGCGTCTTTACAGGCTC	3420
Db	3361		3420
QY	3421	ACCACATAGACGCCATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTAC	3480
Db	3421		3480
QY	3481	CTGCTAGCATACAGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACTCCATCGTGGAC	3540
Db	3481		3540
QY	3541	CAAAATGTGAAGTGTCTCATACGGCTAAAGCTACGCTGACGGGCAACGCGCCCTGCTG	3600
Db	3541		3600
QY	3601	TATAGCTGGAGCGGTTCAAAACGAGGTTTACTACCAACACACCGCCATTAACCAATACATC	3660
Db	3601		3660
QY	3661	ATGGCATGATGTTCGGCTGACCTGGAGTGTGTGACAGACCTGGGTGCTGTAGGCGGA	3720
Db	3661		3720
QY	3721	GTCTTAGCAGCTCTGGCGCGTATTTGCTGACAAACAGGACGCTGTGTCATTGTGGGCGAG	3780
Db	3721		3780
QY	3781	ATCATCTTGTCCGAAAGCGGCGCATCATTTCCGACAGGAAAGTCTTTTACCGGGAGTTC	3840
Db	3781		3840
QY	3841	GATCAGATGAAAGTGGCGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTCGCC	3900

Db	3841	GATGAGATGAAAGATGCGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTGCC	3900
QY	3901	GAACAATTTCAACAGAGCAATCGGTTCTGCAAAACAGCCCAACCAAGCAGGAGGCT	3960
Db	3901		3960
QY	3961	GCTGCTCCCGTGTGGAATCCAAAGTGGCGGACCTCGAAGCCCTTCTGGCGAAGCATATG	4020
Db	3961		4020
QY	4021	TGCAATTTTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTGTCGCTGCAACCCC	4080
Db	4021		4080
QY	4081	GCGATAGCATCACTGATGCAATTCACAGCCCTTATCACCAGCCCGCTCACACCCCAACAT	4140
Db	4081		4140
QY	4141	ACCTCTCTGTTTAAATCTCTGCGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGGCT	4200
Db	4141		4200
QY	4201	GCTTCTGCTTTCGTAGCGCGCATCGCTGGAGCGGCTGTTGCGAGCATAGGCTTGGG	4260
Db	4201		4260
QY	4261	AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTCGTGCC	4320
Db	4261		4320
QY	4321	TTTAAAGTCAAGCGGCGAGATGCCCTCAACGAGCACTGGTTAACTTACCTTCTGCT	4380
Db	4321		4380
QY	4381	ATCCTCTCCCTCGCGCCCTAGTGTGGGGTGTGTGCGAGCGATACCTGCGTGGCAC	4440
Db	4381		4440
QY	4441	GTGGGCCCGAGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCTTGGCTTGGCGG	4500
Db	4441		4500
QY	4501	GGTAAACAGCTTCCCGCACGACTATGCTGTAGAGGAGCGCTGAGCACGCTGTCAC	4560
Db	4501		4560
QY	4561	CAGATCTCTTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTACCAGTGGATCAAC	4620
Db	4561		4620
QY	4621	GAGGACTGTCTCACGCGCATGCTCCGGCTGTGGCTAAGAGATGTTTGGGATTTGGATATGC	4680
Db	4621		4680
QY	4681	ACGGTGTGTGATTTCAAGACCTGGCTCAAGTCAAGCTCTGCGCGGATTTGGCGGGA	4740
Db	4681		4740
QY	4741	GTCCCTCTCTCATGTCAACGTTGGGTACAGGGAGTCTGGCGGGCGACGCGCATATG	4800
Db	4741		4800
QY	4801	CAAAACCACTGCCATGTGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAGG	4860
Db	4801		4860
QY	4861	ATCGTGGGCGCTAGGACCTGTAGTAAACGCTGATGGAACATTTCCCAATTACCGGTAC	4920
Db	4861		4920
QY	4921	ACCAACGGGCGCTGACGCGCTCCCGGCCCAATTTATCTAGGGCGCTGTGGCGGGTG	4980



```
QY 7141 GACCCACACACCCCTTGGCGGGCTGGTGGGACAGCTAGACACACTCCAGTCAAT 7200
|||
Db 7141 GACCCACACACCCCTTGGCGGGCTGGTGGGACAGCTAGACACACTCCAGTCAAT 7200
|||
QY 7201 TCTGCTAGGCAACATCATCATGTATGCGCCACCTTGGGCAAGGATGATCTCTGATG 7260
|||
Db 7201 TCTGCTAGGCAACATCATCATGTATGCGCCACCTTGGGCAAGGATGATCTCTGATG 7260
|||
QY 7261 ACTCATTTCTTCCATCTCTTCTAGCTCAGGAAACAACTGAAAGACCCCTAGATTCTCAG 7320
|||
Db 7261 ACTCATTTCTTCCATCTCTTCTAGCTCAGGAAACAACTGAAAGACCCCTAGATTCTCAG 7320
|||
QY 7321 ACTAGGGCGCTTACTTCCATTTAGGCACTTACCTACCTCAGATCATTTCAACAGCTC 7380
|||
Db 7321 ACTAGGGCGCTTACTTCCATTTAGGCACTTACCTACCTCAGATCATTTCAACAGCTC 7380
|||
QY 7381 CATGGCTTAGCGCATTTTCACTTCCATAGTATCTCTCCAGGTGAGATCAATAGGCTGGCT 7440
|||
Db 7381 CATGGCTTAGCGCATTTTCACTTCCATAGTATCTCTCCAGGTGAGATCAATAGGCTGGCT 7440
|||
QY 7441 TCATGCTCAGGAAACTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAACT 7500
|||
Db 7441 TCATGCTCAGGAAACTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAACT 7500
|||
QY 7501 GTCGGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGCAAGTACCTCTTC 7560
|||
Db 7501 GTCGGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGCAAGTACCTCTTC 7560
|||
QY 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCATCCCGCTCGTCCAGTTGGAT 7620
|||
Db 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCATCCCGCTCGTCCAGTTGGAT 7620
|||
QY 7621 TTATCCAGCTGGTTCGTGCTGTACAGGGGGAGACATATATACAGCCCTGTCTCGT 7680
|||
Db 7621 TTATCCAGCTGGTTCGTGCTGTACAGGGGGAGACATATATACAGCCCTGTCTCGT 7680
|||
QY 7681 GCCGACCCGCTGTTCTGCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7740
|||
Db 7681 GCCGACCCGCTGTTCTGCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7740
|||
QY 7741 CTACTCCCAACCGATGAAGCGGACCTAACTCACTCCAGGCCAATAGGCCATCTCTCTCT 7800
|||
Db 7741 CTACTCCCAACCGATGAAGCGGACCTAACTCACTCCAGGCCAATAGGCCATCTCTCTCT 7800
|||
QY 7801 TTTCCCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 7860
|||
Db 7801 TTTCCCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 7860
|||
QY 7861 TTTTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 7920
|||
Db 7861 TTTTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 7920
|||
QY 7921 TAGCTGGAAGGCTCCGTGAGCGGCTTGAAGTGGAGAGTGTCTGATCTGGCCCTCTCTGC 7980
|||
Db 7921 TAGCTGGAAGGCTCCGTGAGCGGCTTGAAGTGGAGAGTGTCTGATCTGGCCCTCTCTGC 7980
|||
QY 7981 AGATCAAGT 7989
|||
Db 7981 AGATCAAGT 7989
|||

RESULT 3
SSE242652
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SSE242652
Hepatitis C virus replicon I377/NS3-3'UTR.
AJ242652
AJ242652.1 GI:5441834
core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3
proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B
gene; NS4B protein; NS5A gene; NS5A phosphoprotein; NS5B gene; NS5B
RNA dependant RNA polymerase; polyprotein.
Hepatitis C virus replicon I377/NS3-3'UTR
Hepatitis C virus replicon I377/NS3-3'UTR
```

```
artificial sequences; vectors.
1
Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and
Bartenschlager,R.
Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
line
Science 285 (5424), 110-113 (1999)
99322193
MEDLINE
10390360
PUBMED
2 (bases 1 to 7989)
Bartenschlager,R.
Direct Submission
Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,
Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
55131 Mainz, GERMANY
FEATURES
Location/Qualifiers
1..7989
/organism="Hepatitis C virus replicon I377/NS3-3'UTR"
/mol_type="other RNA"
/db_xref="taxon:95363"
/focus
/notes="Neomycin-selectable bicistronic subgenomic
hepatitis C virus replicon"
join(1..341,1801..7758)
/organism="Hepatitis C virus"
/mol_type="other RNA"
/strain="1b"
/isolate="Conl"
/db_xref="taxon:11103"
/tissue_type="liver"
1190..1800
/organism="Encephalomyocarditis virus"
/mol_type="other RNA"
/db_xref="taxon:12104"
1..341
/evidence="experimental"
342..1181
/genes="core-neo"
342..1181
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46912.1"
/db_xref="GI:5441835"
/translation="MSTNPKPQKTKGRAMIEQDGLHAGSPAAWVERLFQYDWAQQT
GCSDAVFLSAQGRVFLVKTLDSGALNELODEAARLSWLATGVPFCAVLVDVTEA
GRDWLLGVPGQDLLSSHLAPAEKVSINADAMRRLHLDLPATCPDPAKRRIERAR
TRMEAGLVDDDLDEHQGLAPAEKPARLKARMPDGEDLVVTHGDACLNNIMVNGRF
SGFIDCRLGVADRYQDIALATRDIAELGGEWADRFVLVYIAAPDSQRIAFYRLLD
EPF"
misc_feature
1190..1800
/notes="internal ribosome entry site (IRES)"
1801..7758
/codon_start=1
/product="non-structural polyprotein"
/protein_id="CAB46913.1"
/db_xref="GI:5441836"
/translation="MAPITAYSQQTGRLGGLCIITSLGRDRNQVEGEVQVSTATQSF
LATVNGVCTVYVHAGSKTLAGPKGPTQMTVNTVDLGVQAPGARSLTPTCTGCS
SDLYLTVRHADVI PVERRGDSGLLSPREVSVLKGSGGPLLCPSGHAGVIFRAAVC
TRGVAKAVDFPVESNETTMRSPVPTDSSPPAVPOTFOVAHLHAPTSGSKTKVPA
YAAQGGKVDVLIICDECHSTDDITLIGTGLVDQDAETAGRLVVLATATPPGVSVPFH
DNIEVALSGTGEIPYGAKEIETIKGGRHLIFCHSKKCDLAAKLSGLGNVAY
YRGLDVSVPTGSDVIVATDALMTGTGDFSDVIDCNTCTVQTVDFSLDPTIETT
TVPDQAVSRQRGRTGRGMGIYRVTPERPSGMFSDSVLCECCYDAGCAWFLVPA
TVSVRLRVLYNTFGLPQDHLFFWESVETGLTHIDAHFLSQTQKQADNFFVLVYQA
TVCRAQAQPPSWDQMKKILRLHPTPLLYRLGAVQNEVTHITPITIKMACM
SADLEVTVTVLVGGLAAALAAAYCLTGGSVVIGKRIILSGKPAIIPDREVLFRFDE
MEBCASHLPYIEQMQLAEQKQKALGLQTATKQAAAPVSVESKWRLEAFWAKH
WNFISGIOVLGLSTLPGNPAJASLMAFTASITPLTQHTLLFNILGWWAAQLAPP
SAASAFVGAGIAGAAVSGISGLGKVLVDILAGYGAGVAGALVAFKVMSEMPSTEDLVN
```

LLPALLSPALVGVVCAAILRRHVGPGEVQVMMNRLIAFASRGNHVSPTHVPESD  
AAARVTOILSSITLQILKRLHOWINEDCSTPCSGWLRDWDWMLCTVLTOPKTLQOS  
KLPLRLVPRESQGVYKWRGIMQITPCQAQITGHVKNSGMRVLPRTCSNT  
WHGTFPINAITYTGPCTPAPNYSRALMRVABEYEVTRVGDPHIVTVGMTIDNKKCP  
CQVPAPEFTEVDGRLHRYAPACPLUREEVTFLVNLQYLVDGSQLCEPDPVAVL  
TSMLEDPSHI TAETAKRLARLSPSLASASQLSAPSLKATCTRHDSPPADLIEA  
NLLMRQMGKNTIRVESNKVILDSFELPQAEDEREVSPAEILRRSRKPPRAMPI  
WARPDYNPLLESMDPDVVPVHVHCPPLPAKAPPIPPRRKRTVVLSESTVSALA  
ELATVTFSSSESSAVDSGTAPSDPDSDGDGSDVESYSSMPPLEGEPPDPLSDG  
SWSVTSBEASEDVVCCSGSYWTGALI TPCAABETKLPINALNSLLRHNDLVYATTS  
RSASLURKKVTFDRLOVDDHVDVLMKEMKASTVKAJLLSVEACKLTPPHSARSK  
PYGKADVRNYSKKAHNIRSVKOLLEDETPTDITIMAKNEVFCVQPEKGRKPAR  
LIVFDPLGVNICEKMAJYVSTLPQAVMGSSYGFQSPGQVRELVMNAKAKCPMG  
FAYDRCFDSTWENDI VESIVYOCCLAPAROAIRSLTERLVIGGLPLTNSKQNC  
GYRRCASGLVTSQNTLTCLYKAAACRAKALQDCTMLVCGDPLVIVCESAGTQED  
EASLRAFTEAMTRYISAPGDPDPKPEYDLELITSNSNVSVHADSGKRVYILTRPDT  
PLAARWETARHTPVNSWIGNIIMYAPTLWAMILMTHFFSILLAOEYLKALDCQIY  
GACYSIEFLDPAQIIFORLHGLSFAFSHSYSGEINRVASCLRLKGLVPLRVWRHARS  
VRALLSQGGRAATCGKLVFNWAVRTKLKLTPIPAASQLDLSWFEVAGYSGGDI VHSI  
SRARPRMFMWCLLLLSVGVGIYLLPNR"

mat\_peptide

1801..3696

/gene="NS3"

/product="NS3 proteinase/helicase"

mat\_peptide

3697..3858

/gene="NS4B"

/product="NS3/4A proteinase cofactor"

mat\_peptide

3859..4641

/gene="NS4B"

/product="NS4B protein"

mat\_peptide

4642..5982

/gene="NS5A"

/product="NS5A phosphoprotein"

mat\_peptide

5983..7755

/gene="NS5B"

/product="NS5B RNA dependant RNA polymerase"

gene

1801..3696

/gene="NS3"

gene

3697..3858

/gene="NS4A"

gene

3859..4641

/gene="NS4B"

gene

4642..5982

/gene="NS5A"

gene

5983..7755

/gene="NS5B"

7759..7989

Query Match 100.0%; Score 7987.4; DB 12; Length 7989;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCATTTGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60

DB 1 GCCAGCCCCCATTTGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60

QY 61 TCTTCACGACAGAAAGCGTCTAGCCATGCGTTAGTATGATGATGTCGTGAGCTCCAGGAC 120

DB 61 TCTTCACGACAGAAAGCGTCTAGCCATGCGTTAGTATGATGTCGTGAGCTCCAGGAC 120

QY 121 CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATGCCAG 180

DB 121 CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATGCCAG 180

QY 181 GACGACCGGGTCCCTTCTTGTGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC 240

DB 181 GACGACCGGGTCCCTTCTTGTGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC 240

QY 241 CGGAGCTGTAGCCGAGTGTGTTGGGTGCGAAAGGCCCTTGTGGTACTGCTGATAGG 300

DB 241 GCGAGACTGTAGCCGAGTGTGTTGGGTGCGAAAGGCCCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGAGTGCCCGGGAGGCTCTCGTAGACCGTGCACCATGAGCAGGAATCCTAAAC 360

DB 301 GTGCTTGAGTGCCCGGGAGGCTCTCGTAGACCGTGCACCATGAGCAGGAATCCTAAAC 360

QY 361 CTCAAAGAAAACCAAAAGGCGCGCATGATTGAAACAAGATGATGCAAGCAGGTTCTC 420

DB 361 CTCAAAGAAAACCAAAAGGCGCGCATGATTGAAACAAGATGATGCAAGCAGGTTCTC 420

QY 421 CGGCGCGCTTGGTGGAGAGGCTATTTCGGCTATGATGCGGCAACAAGCAATCGGCTGCT 480

DB 421 CGGCGCGCTTGGTGGAGAGGCTATTTCGGCTATGATGCGGCAACAAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGGCGGCGCGCGGCTCTTTTGTCAAGACCG 540

DB 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGGCGGCGCGCGGCTCTTTTGTCAAGACCG 540

QY 541 ACCTTTCGGTGGTCCCTGTAATGAATGCGAGGAGCGCGGCTATCGTGGCTGGCCA 600

DB 541 ACCTTTCGGTGGTCCCTGTAATGAATGCGAGGAGCGCGGCTATCGTGGCTGGCCA 600

QY 601 CGACGGGCGTTCCTTTCGCGAGCTGTGCTGCGAGTGTGCTCACTGAAGCGGAAGGACTGC 660

DB 601 CGACGGGCGTTCCTTTCGCGAGCTGTGCTGCGAGTGTGCTCACTGAAGCGGAAGGACTGC 660

QY 661 TGCTATTGGGCGAAGTGCAGGAGGAGGATCTCTGCTCATCTCACCTTGTCTCTGCGAGA 720

DB 661 TGCTATTGGGCGAAGTGCAGGAGGAGGATCTCTGCTCATCTCACCTTGTCTCTGCGAGA 720

QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

DB 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 CATTCGACCAACCAAGCGAATCATGCTGAGGAGGAGGATCTGCTGCTGCTGCTGCTGCTGCT 840

DB 781 CATTCGACCAACCAAGCGAATCATGCTGAGGAGGAGGATCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 TTGTCGATCAGGATGATCTGACGAGGAGGATCAGGGGCTGCGGCGAGCGAATCTGCTGCTG 900

DB 841 TTGTCGATCAGGATGATCTGACGAGGAGGATCAGGGGCTGCGGCGAGCGAATCTGCTGCTG 900

QY 901 CAGGCTCAAGGCGCGCATGCGGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

DB 901 CAGGCTCAAGGCGCGCATGCGGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

QY 961 GCTTCCGCAATATCATGTTGGAATAATGCGGCTTTCTGAGTATCATGCTGCTGCTGCTGCTGCT 1020

DB 961 GCTTCCGCAATATCATGTTGGAATAATGCGGCTTTCTGAGTATCATGCTGCTGCTGCTGCTGCT 1020

QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGGTTGGCTACCGGTGATATTGCTGAAGAGC 1080

DB 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGGTTGGCTACCGGTGATATTGCTGAAGAGC 1080

QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTGCTGCTTTCAGGATATCGCGCTGCTGCTGCTGCTGCT 1140

DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTGCTGCTTTCAGGATATCGCGCTGCTGCTGCTGCTGCT 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTATTAACAGACCAACAG 1200

DB 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTATTAACAGACCAACAG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAAGCTTACTGCGC 1260

DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAAGCTTACTGCGC 1260

QY 1261 CGAAGCGCTTGGAAATAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

DB 1261 CGAAGCGCTTGGAAATAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTTCTTCTGACGAGCATTCCT 1380

DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTTCTTCTGACGAGCATTCCT 1380

QY 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTTGATGCTGCTGAGGAAGCA 1440



Db 1381 AGGGGCTTTTCCCTCTCGCCCAAGGATGCAAGGCTGTGTGAATCTGTGTAAGGAGCA 1440  
QY 1441 GTTCTCTTGAAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCTTTGCGAGCGGG 1500  
Db 1441 GTTCTCTTGAAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCTTTGCGAGCGGG 1500  
QY 1501 AACCCCCCACTCGGACAGGTGCTCTGGGCCAAAGCCAGCTGTATAGATACACCT 1560  
Db 1501 AACCCCCCACTCGGACAGGTGCTCTGGGCCAAAGCCAGCTGTATAGATACACCT 1560  
QY 1561 GCAAGCGGCAACAAACCCAGTCCAGTTGTGAGTTGAGTAGTTGTGAAAGAGTCAAA 1620  
Db 1561 GCAAGCGGCAACAAACCCAGTCCAGTTGTGAGTTGAGTAGTTGTGAAAGAGTCAAA 1620  
QY 1621 TGCGTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
Db 1621 TGCGTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTACATGTTTGTAGTCGAGTTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTACATGTTTGTAGTCGAGTTAAAA 1740  
QY 1741 AAGCTCTAGGCCCCCGCAACACGGGACGTGTTTTTCCCTTTGAAAAACACGATATACC 1800  
Db 1741 AAGCTCTAGGCCCCCGCAACACGGGACGTGTTTTTCCCTTTGAAAAACACGATATACC 1800  
QY 1801 ATGGCGCTATTAGCGCTTACTCCCAACAGCGGAGGCTTCTGCGTGCATCATCACT 1860  
Db 1801 ATGGCGCTATTAGCGCTTACTCCCAACAGCGGAGGCTTCTGCGTGCATCATCACT 1860  
QY 1861 AGCTCACAGCGCGGACAGGACAGGTGAGGGGAGTCCAAAGTGGTCTCCACGCA 1920  
Db 1861 AGCTCACAGCGCGGACAGGACAGGTGAGGGGAGTCCAAAGTGGTCTCCACGCA 1920  
QY 1921 ACACAATCTTCTCGCGACTCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
Db 1921 ACACAATCTTCTCGCGACTCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
QY 1981 GGCTAAAGACCTTTCGGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 1981 GGCTAAAGACCTTTCGGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACCTGTGGCTGGCAAGCGCCCGGGCGGCTTCTTGACACCATGCCTGC 2100  
Db 2041 CAGGACCTGTGGCTGGCAAGCGCCCGGGCGGCTTCTTGACACCATGCCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCCGATGTCATTCGGTGGCGGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCCGATGTCATTCGGTGGCGGCGG 2160  
QY 2161 GGCAGCAGCAGGGGAGCTTCTCTCCCGAGGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
Db 2161 GGCAGCAGCAGGGGAGCTTCTCTCCCGAGGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGCTCACTGTCTGCCCTCGGGCAGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
Db 2221 GCGGCTCACTGTCTGCCCTCGGGCAGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCGGTGACCGAGACATTCAGGGT 2400  
Db 2341 CGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCGGTGACCGAGACATTCAGGGT 2400  
QY 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGACACTAAGTGGCGGTGGTAGCA 2460  
Db 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGACACTAAGTGGCGGTGGTAGCA 2460  
QY 2461 GCCCAAGGGGTATAGGTGCTTGTCTGAAACCCGTCCGTCGCGCCACCTAGTTCGGG 2520  
Db 2461 GCCCAAGGGGTATAGGTGCTTGTCTGAAACCCGTCCGTCGCGCCACCTAGTTCGGG 2520

QY 2521 GCGTATATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGGGTAAAGCACCATC 2580  
Db 2521 GCGTATATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGGGTAAAGCACCATC 2580  
QY 2581 ACCACGGGTGCCCCATCAGCTACTCACCCTATGGCAAGTTTCTTGCACGCGTGGTTGC 2640  
Db 2581 ACCACGGGTGCCCCATCAGCTACTCACCCTATGGCAAGTTTCTTGCACGCGTGGTTGC 2640  
QY 2641 TCTGGGGGCGCTTATGACATCATATATGTGATGTGATGCCACTCAACTGACTCGACCAT 2700  
Db 2641 TCTGGGGGCGCTTATGACATCATATATGTGATGTGATGCCACTCAACTGACTCGACCAT 2700  
QY 2701 ATCTCTGGGATCGGACAGTCTTGGACCAAGCGGACAGCGCTGGAGCGGACTCGCTG 2760  
Db 2701 ATCTCTGGGATCGGACAGTCTTGGACCAAGCGGACAGCGCTGGAGCGGACTCGCTG 2760  
QY 2761 CTGCCCACCGCTACGCTCGGGATCGGTCAACCGTGCCACATCAACATCGAGGAGGTG 2820  
Db 2761 CTGCCCACCGCTACGCTCGGGATCGGTCAACCGTGCCACATCAACATCGAGGAGGTG 2820  
QY 2821 GCTCTCTCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTCTCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACACCTCATTTTCTGCCATTCCAAAGAAATGTGATGAGCTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACACCTCATTTTCTGCCATTCCAAAGAAATGTGATGAGCTCGCGCG 2940  
QY 2941 AAGCTGTGCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATGATCGCTC 3000  
Db 2941 AAGCTGTGCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATGATCGCTC 3000  
QY 3001 ATACCAACTAGCGGACAGCTCATTTGTCTGAGCAACGAGCTCTAATAGCGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGGACAGCTCATTTGTCTGAGCAACGAGCTCTAATAGCGGCTTTACC 3060  
QY 3061 GCGGATTTGACTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
Db 3061 GCGGATTTGACTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
QY 3121 CTGACCCGACCTTCAACCTTGTAGAGACGACCGCTGCAAGACGCGGTGTCAACGCTCG 3180  
Db 3121 CTGACCCGACCTTCAACCTTGTAGAGACGACCGCTGCAAGACGCGGTGTCAACGCTCG 3180  
QY 3181 CAGCGCGGAGGACAGTGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGCGGAGGACAGTGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTGCTGCTGCGAGTGTATGACGCGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTGCTGCTGCGAGTGTATGACGCGGGCTGT 3300  
QY 3301 GCTTGTGTACGAGCTCACGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTGTACGAGCTCACGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
Db 3361 CCAGGTTGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTTAAGCAGCAGGACACACTTCCCTTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTTAAGCAGCAGGACACACTTCCCTTAC 3480  
QY 3481 CTGCTAGTACACGAGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCACTCGTGGAC 3540  
Db 3481 CTGCTAGTACACGAGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCACTCGTGGAC 3540  
QY 3541 CAAATGTGAAGTGTCTCATACGCTTAAAGCTTACGCTGCAAGGGCCAAACGCGCTGCTG 3600  
Db 3541 CAAATGTGAAGTGTCTCATACGCTTAAAGCTTACGCTGCAAGGGCCAAACGCGCTGCTG 3600



3601 TATAGCTGGAGCGCTTCAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
Db  
3601 TATAGCTGGAGCGCTTCAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
QY  
3661 ATGGCANTGATGTGGCTGACCTGAGAGTGGTCAAGAGCACCTGGTGTGGTGGG 3720  
Db  
3661 ATGGCANTGATGTGGCTGACCTGAGAGTGGTCAAGAGCACCTGGTGTGGTGGG 3720  
QY  
3721 GTCTAGACAGCTGTGGCGCGTATTCCTGCAACAGGAGCGTGGTCAATGTGGCAGG 3780  
Db  
3721 GTCTAGACAGCTGTGGCGCGTATTCCTGCAACAGGAGCGTGGTCAATGTGGCAGG 3780  
QY  
3781 ATCACTCTTCGGAAAGCGGCCATCATCTCCCGACAGAGGAAGTCTTTACCGGAGTTC 3840  
Db  
3781 ATCACTCTTCGGAAAGCGGCCATCATCTCCCGACAGAGGAAGTCTTTACCGGAGTTC 3840  
QY  
3841 GATGAGATGGAAGTGGCCCTCAACCTCCCTTATCATCGAAACAGGGAATGACGCTCGCC 3900  
Db  
3841 GATGAGATGGAAGTGGCCCTCAACCTCCCTTATCATCGAAACAGGGAATGACGCTCGCC 3900  
QY  
3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGAGGCT 3960  
Db  
3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGAGGCT 3960  
QY  
3961 GCTGCTCCGCTGGTGAATPCCAAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
Db  
3961 GCTGCTCCGCTGGTGAATPCCAAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
QY  
4021 TGGAAATTCATCAGCGGGATCAATATTTAGCAGGCTTGTCACATCTGCTGGCAACCCC 4080  
Db  
4021 TGGAAATTCATCAGCGGGATCAATATTTAGCAGGCTTGTCACATCTGCTGGCAACCCC 4080  
QY  
4081 GGGATAGCATCATGATGSCATTCACAGCTCTATCACAGCCCGCTCACACCCCAACAT 4140  
Db  
4081 GGGATAGCATCATGATGSCATTCACAGCTCTATCACAGCCCGCTCACACCCCAACAT 4140  
QY  
4141 ACCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGCTCTCCACGCGCT 4200  
Db  
4141 ACCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGCTCTCCACGCGCT 4200  
QY  
4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTGGG 4260  
Db  
4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTGGG 4260  
QY  
4261 AAGTGTCTTGTGATATTTGGCAGTTATGGAGCAGGGGTGGCAGGCGCGCTCGTGCC 4320  
Db  
4261 AAGTGTCTTGTGATATTTGGCAGTTATGGAGCAGGGGTGGCAGGCGCGCTCGTGCC 4320  
QY  
4321 TTAAAGGTATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
Db  
4321 TTAAAGGTATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
QY  
4381 ATCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCGAGCATCTGGCTCGGCAC 4440  
Db  
4381 ATCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCGAGCATCTGGCTCGGCAC 4440  
QY  
4441 GTGGCCCAAGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
Db  
4441 GTGGCCCAAGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
QY  
4501 GGTAACACGCTCTCCCAACGACATATGTGCTGAGAGCGAGCTGACAGCATGTCAC 4560  
Db  
4501 GGTAACACGCTCTCCCAACGACATATGTGCTGAGAGCGAGCTGACAGCATGTCAC 4560  
QY  
4561 CAGATCTCTTAGTCTTTACCATCACTCAGCTCTGAAAGAGCTTCAACAGTGGATCAAC 4620  
Db  
4561 CAGATCTCTTAGTCTTTACCATCACTCAGCTCTGAAAGAGCTTCAACAGTGGATCAAC 4620  
QY  
4621 GAGGACTGTCTCACGCCATGTCCGGCTCGTGGCTTAAGAGATGTTTGGATTTGATATGC 4680  
Db  
4621 GAGGACTGTCTCACGCCATGTCCGGCTCGTGGCTTAAGAGATGTTTGGATTTGATATGC 4680  
QY  
4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTCGCGGATGCGGG 4740

4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCTGCGGATGCGCGGA 4740  
QY  
4741 GTCCCTCTTCTCATGTCAACGTGGGTAAAGGGAGTCTGGCGGGCGACGSCATCATG 4800  
Db  
4741 GTCCCTCTTCTCATGTCAACGTGGGTAAAGGGAGTCTGGCGGGCGACGSCATCATG 4800  
QY  
4801 CAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAATAACCGTTCCATGAGG 4860  
Db  
4801 CAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAATAACCGTTCCATGAGG 4860  
QY  
4861 ATCTGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTCGCCATTAACGCGTAC 4920  
Db  
4861 ATCTGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTCGCCATTAACGCGTAC 4920  
QY  
4921 ACCACGGGCCCCGTGACGCCCCCTCCCGCGCCAAATTAATCTAGGCGCTGTGGGGGCTG 4980  
Db  
4921 ACCACGGGCCCCGTGACGCCCCCTCCCGCGCCAAATTAATCTAGGCGCTGTGGGGGCTG 4980  
QY  
4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGGCTG 5040  
Db  
4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGGCTG 5040  
QY  
5041 ACCACTGAACAGTAAAGTGCCTGTGAGTTCCGGCCCCCGAAATTTCTCACAGAAGTG 5100  
Db  
5041 ACCACTGAACAGTAAAGTGCCTGTGAGTTCCGGCCCCCGAAATTTCTCACAGAAGTG 5100  
QY  
5101 GATGGGTGCGGTTGCAACAGTACGCTCCAGGTGCAAAACCCCTCTACGGGAGAGGTC 5160  
Db  
5101 GATGGGTGCGGTTGCAACAGTACGCTCCAGGTGCAAAACCCCTCTACGGGAGAGGTC 5160  
QY  
5161 ACATTCCTGTGCGGCTCAATCAATACCTGTTGGGTGACAGTCCCATGCCAGCCGAA 5220  
Db  
5161 ACATTCCTGTGCGGCTCAATCAATACCTGTTGGGTGACAGTCCCATGCCAGCCGAA 5220  
QY  
5221 CGGACGTAGCAGTCTCACTTCCATGTCTACCGACCCCTCCACATACGGCGGAGCG 5280  
Db  
5221 CGGACGTAGCAGTCTCACTTCCATGTCTACCGACCCCTCCACATACGGCGGAGCG 5280  
QY  
5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
Db  
5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
QY  
5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCCGTGATGCTCCCGGACGCTGAC 5400  
Db  
5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCCGTGATGCTCCCGGACGCTGAC 5400  
QY  
5401 CTCATCGAGGCAACCTCTCTGCGGCGAGAGATGGCGGGAAACATCACCCGCGTGGAG 5460  
Db  
5401 CTCATCGAGGCAACCTCTCTGCGGCGAGAGATGGCGGGAAACATCACCCGCGTGGAG 5460  
QY  
5461 TCAGAAATATAGGTAGTAAATTTTGAATCTTTTCAGCGCTCCAAAGCGGAGGAGATGAG 5520  
Db  
5461 TCAGAAATATAGGTAGTAAATTTTGAATCTTTTCAGCGCTCCAAAGCGGAGGAGATGAG 5520  
QY  
5521 AGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGGAATTCCTCTGACGATG 5580  
Db  
5521 AGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGGAATTCCTCTGACGATG 5580  
QY  
5581 CCCATATGGGCAACCGCGGATTAACACCCCTCCACTGTTAGAGTCTGGAAGGACCCGGAC 5640  
Db  
5581 CCCATATGGGCAACCGCGGATTAACACCCCTCCACTGTTAGAGTCTGGAAGGACCCGGAC 5640  
QY  
5641 TAGTCCCTCAGTGGTACAGGGGTGCCATTCGCGCTGCAAGGCCCCCTCCCGATACA 5700  
Db  
5641 TAGTCCCTCAGTGGTACAGGGGTGCCATTCGCGCTGCAAGGCCCCCTCCCGATACA 5700  
QY  
5701 CCTCCAGGAGAGAGGACGTTGTCTGTGACAACTACCGTGTCTTCTGCTTGGG 5760  
Db  
5701 CCTCCAGGAGAGAGGACGTTGTCTGTGACAACTACCGTGTCTTCTGCTTGGG 5760  
QY  
5761 GAGTCTGCGCAAAAGACCTTCGCGAGCTCCGAATGCTCGGCGCTCGACAGCGGACGGA 5820

5761 GAGTCCGCCAAGAACCTTGGCAGCTCCGAAATGCTGGCCGTCGACAGCGGCACGGCA 5820 Db  
5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGACCGGGATCCGACGTTGAGTCGTAC 5880 Qy  
5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGACCGGGATCCGACGTTGAGTCGTAC 5880 Db  
5881 TCTCCATGCCCCCTTGAAGGGGAGCGGGGATCCGATCTCAGACAGCGGTTCTGG 5940 Qy  
5881 TCTCCATGCCCCCTTGAAGGGGAGCGGGGATCCGATCTCAGACAGCGGTTCTGG 5940 Db  
5941 TCTACCGTAAGCAGAGAGCTAGTAGAGACGTCGTCTGCTCGATGTCCTACACATGG 6000 Qy  
5941 TCTACCGTAAGCAGAGAGCTAGTAGAGACGTCGTCTGCTCGATGTCCTACACATGG 6000 Db  
6001 ACAGGCGCCTGATCAGCCATGCGCTGCGGAGAAACAAAGTCCGCTCAATGCACTG 6060 Qy  
6001 ACAGGCGCCTGATCAGCCATGCGCTGCGGAGAAACAAAGTCCGCTCAATGCACTG 6060 Db  
6061 AGCAACTCTTGTCCGTCACCAACACTTGGTCTATGCTACACATCTCGCAGCGCAAGC 6120 Qy  
6061 AGCAACTCTTGTCCGTCACCAACACTTGGTCTATGCTACACATCTCGCAGCGCAAGC 6120 Db  
6121 CTGCGCAGAGAGAGTCACTTTGACAGACTGCAGTCTGACGACCACTACCGGAC 6180 Qy  
6121 CTGCGCAGAGAGAGTCACTTTGACAGACTGCAGTCTGACGACCACTACCGGAC 6180 Db  
6181 GTGCTCAAGAGATGAAGGCGTAAGGCTCCACAGTTAAGCTAAACTTCTATCCGTGGAG 6240 Qy  
6181 GTGCTCAAGAGATGAAGGCGTAAGGCTCCACAGTTAAGCTAAACTTCTATCCGTGGAG 6240 Db  
6241 GAAGCTGTAAAGCTGACGCCCCACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300 Qy  
6241 GAAGCTGTAAAGCTGACGCCCCACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300 Db  
6301 GAGCTCCGAACCTATCCAGCAAGGCGTTAAACACATCCGCTCCGTTGGAAGCACTTG 6360 Qy  
6301 GAGCTCCGNACTATCCAGCAAGGCGTTAAACACATCCGCTCCGTTGGAAGCACTTG 6360 Db  
6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATGAGTTTCTGC 6420 Qy  
6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATGAGTTTCTGC 6420 Db  
6421 GTCCACACAGAGAGGGGCGCAGCAGCTCGCTTATCGATATCCAGATTTGGGG 6480 Qy  
6421 GTCCACACAGAGAGGGGCGCAGCAGCTCGCTTATCGATATCCAGATTTGGGG 6480 Db  
6481 GTTCGTGTGCGAGAAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGGTG 6540 Qy  
6481 GTTCGTGTGCGAGAAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGGTG 6540 Db  
6541 ATGGCTCTTCAACGGAATTCGAATCTCTCCTGGACAGCGGTTCGAGTTCTGTGTAAT 6600 Qy  
6541 ATGGCTCTTCAACGGAATTCGAATCTCTCCTGGACAGCGGTTCGAGTTCTGTGTAAT 6600 Db  
6601 GCCTGGAAGCGAAGAAATGCGCTATGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660 Qy  
6601 GCCTGGAAGCGAAGAAATGCGCTATGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660 Db  
6661 ACAGCTCACTGAGATGACATCCGTTGAGAGTCAATCTACCAATGTTGTGACTGGCC 6720 Qy  
6661 ACAGCTCACTGAGATGACATCCGTTGAGAGTCAATCTACCAATGTTGTGACTGGCC 6720 Db  
6721 CCCGAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGTTTACATCGGGGCGCCCTG 6780 Qy  
6721 CCCGAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGTTTACATCGGGGCGCCCTG 6780 Db  
6781 ACTAAATCTTAAAGGAGAACTGGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840 Qy  
6781 ACTAAATCTTAAAGGAGAACTGGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840 Db  
6841 ACCAGCTGCGGTATACCCCTCAATGTTACTTGAAGCCGCTGCGGCTCTCGAGCTGCG 6900 Qy  
6841 ACCAGCTGCGGTATACCCCTCAATGTTACTTGAAGCCGCTGCGGCTCTCGAGCTGCG 6900 Db

6901 AAGCTCCAGGACTGCAAGATGCTGATATCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960 Qy  
6901 AAGCTCCAGGACTGCAAGATGCTGATATCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960 Db  
6961 GCGGGAGCCCAAGAGACGAGCGAGCTTACGGGCTTTCACGAGGCTATGACTAGATAC 7020 Qy  
6961 GCGGGAGCCCAAGAGACGAGCGAGCTTACGGGCTTTCACGAGGCTATGACTAGATAC 7020 Db  
7021 TCTGCCCTTGGGAGCCGCCAAACAGAAATACGATTTGGAGTTGATAACATCATGC 7080 Qy  
7021 TCTGCCCTTGGGAGCCGCCAAACAGAAATACGATTTGGAGTTGATAACATCATGC 7080 Db  
7081 TCTTCCAAATGTGTCAGTCCGCGACAGTGCATCTGCGCAAAAGGTTGACTATCTCAACCGT 7140 Qy  
7081 TCTTCCAAATGTGTCAGTCCGCGACAGTGCATCTGCGCAAAAGGTTGACTATCTCAACCGT 7140 Db  
7141 GACCCACACACCCCTTCCGCGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200 Qy  
7141 GACCCACACACCCCTTCCGCGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200 Db  
7201 TCTTGGCTAGGCAACATCATATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260 Qy  
7201 TCTTGGCTAGGCAACATCATATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260 Db  
7261 ACTCATTTCTTCCATCTTCTAGCTCAGGACAACTTGAANAAGCCCTAGATTCTCAG 7320 Qy  
7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGAANAAGCCCTAGATTCTCAG 7320 Db  
7321 ATCTAGCGGCTGTGTTACTTCCATGAGCCACTTACCTACCTCAGATCATTTCAACGACTC 7380 Qy  
7321 ATCTAGCGGCTGTGTTACTTCCATGAGCCACTTACCTACCTCAGATCATTTCAACGACTC 7380 Db  
7381 CATGGCTTAGCGCAATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT 7440 Qy  
7381 CATGGCTTAGCGCAATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT 7440 Db  
7441 TCAATGCTCAGGAACTTGGGTACCGCTTGCAGTCTGGAGACATGGGCGCAAGT 7500 Qy  
7441 TCAATGCTCAGGAACTTGGGTACCGCTTGCAGTCTGGAGACATGGGCGCAAGT 7500 Db  
7501 GTCCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560 Qy  
7501 GTCCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560 Db  
7561 AACTGGGCGATAGGACCAAGCTCAAACTCACTCCAAATCCCGCTCGCTCCAGTTGGAT 7620 Qy  
7561 AACTGGGCGATAGGACCAAGCTCAAACTCACTCCAAATCCCGCTCGCTCCAGTTGGAT 7620 Db  
7621 TTAATCAGCTGTTCTGTTGCTGTTACAGGGGGAGACATATATCACAGCTGTCTCGT 7680 Qy  
7621 TTAATCAGCTGTTCTGTTGCTGTTACAGGGGGAGACATATATCACAGCTGTCTCGT 7680 Db  
7681 GCGGACCCCGCTGTTGTTGCTGTTACTCTACTTCTTAGGGGTAGGCACTAT 7740 Qy  
7681 GCGGACCCCGCTGTTGTTGCTGTTACTCTACTTCTTAGGGGTAGGCACTAT 7740 Db  
7741 CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCAATCTGTGTTT 7800 Qy  
7741 CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCAATCTGTGTTT 7800 Db  
7801 TTTCCCTTT 7860 Qy  
7801 TTTCCCTTT 7860 Db  
7861 TTTTCTCTTT 7920 Qy  
7861 TTTTCTCTTT 7920 Db  
7921 TAGCTGTGAAGGTCCGTGAGCGCTTGACTGACAGAGAGTGTGATCTGGCCTCTCTGC 7980 Qy  
7921 TAGCTGTGAAGGTCCGTGAGCGCTTGACTGACAGAGAGTGTGATCTGGCCTCTCTGC 7980 Db

```
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 4
LOCUS AX739972 10690 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 3 from Patent WO20059321.
ACCESSION AX739972
VERSION AX739972.1 GI:30519246
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS de Francesco, R., Migliaccio, G. and Paonessa, G.
TITLE Hepatitis C virus replicons and replicon enhanced cells
JOURNAL Patent: WO 02059321-A 3 01-AUG-2002;
INSTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.
(IT)

FEATURES
Location/Qualifiers
source 1..10690
/mol_type="unidentified"
/db_xref="taxon:32644"
/note="pHCVNeo.17 coding"

ORIGIN
Query Match 100.0%; Score 7987.4; DB 6; Length 10690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTGGGGGACACCTCCACCATAGATCACTCCCTGTGAGAACTACTG 60
Db 1 GCCAGCCCCGATTGGGGGACACCTCCACCATAGATCACTCCCTGTGAGAACTACTG 60

QY 61 TCTTCACGAGAAAGCGCTAGCCATCGCTAGTATGAGTGTCTGTGAGCCTCCAGGAC 120
Db 61 TCTTCACGAGAAAGCGCTAGCCATCGCTAGTATGAGTGTCTGTGAGCCTCCAGGAC 120

QY 121 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
Db 121 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180

QY 181 GACGACCGGTCCTTTCTTTGATCAACCGCTCAATGCTGAGATTGGGGTCCCGCC 240
Db 181 GACGACCGGTCCTTTCTTTGATCAACCGCTCAATGCTGAGATTGGGGTCCCGCC 240

QY 241 GCGAGACTGTAGCGGAGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCGGAGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGCAGTCCCGGGAGGTCTCGTAGACCGTGACCATGACGACGAACTCTAAAC 360
Db 301 GTGCTTGCAGTCCCGGGAGGTCTCGTAGACCGTGACCATGACGACGAACTCTAAAC 360

QY 361 CTCAAGAAAAACAAAGGCGCGCATGATTGAACAAAGATGGATTGCAACGAGTCTC 420
Db 361 CTCAAGAAAAACAAAGGCGCGCATGATTGAACAAAGATGGATTGCAACGAGTCTC 420

QY 421 CGGCGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAGCAATCGGCTGCT 480
Db 421 CGGCGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGGTGTTCCGGCTGTACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540
Db 481 CTGATGCGCGGTGTTCCGGCTGTACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGGTCCCTGAAATGAATCACTGACGAGGAGCGCGGTATCGTGGCTGGCCA 600
Db 541 ACCTGTCCGGTCCCTGAAATGAATCACTGACGAGGAGCGCGGTATCGTGGCTGGCCA 600

QY 601 CGACGGGCGTTCTTTGGCGCAGCTGTCTGCGAGTGTCTCACTGAACGGGAAAGGACTGGC 660
Db 601 CGACGGGCGTTCTTTGGCGCAGCTGTCTGCGAGTGTCTCACTGAACGGGAAAGGACTGGC 660

QY 661 TGCTATTGGGCGAAGTGCCTGGGCGAGGATCTCTCTCATCTTCACTTGTCTCTGCGGAGA 720
Db 661 TGCTATTGGGCGAAGTGCCTGGGCGAGGATCTCTCTCATCTTCACTTGTCTCTGCGGAGA 720

QY 721 AAGTATCATCATGTGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCCGGTACTCGC 780
Db 721 AAGTATCATCATGTGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCCGGTACTCGC 780

QY 781 CATTCGACCAACCAAGCGAAACATCGCATCGACGAGCAGTACTCGATGGAAGCGGTC 840
Db 781 CATTCGACCAACCAAGCGAAACATCGCATCGACGAGCAGTACTCGATGGAAGCGGTC 840

QY 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGCTCGCGCCAGCCGAACCTGTTGG 900
Db 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGCTCGCGCCAGCCGAACCTGTTGG 900

QY 901 CCAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTGTGACCATGCGGATGCGT 960
Db 901 CCAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTGTGACCATGCGGATGCGT 960

QY 961 GCTTCCCGAATATCATGTGGAATAATGCCGCTTTTCTGGATTCTATCGACTGTGGCGCGC 1020
Db 961 GCTTCCCGAATATCATGTGGAATAATGCCGCTTTTCTGGATTCTATCGACTGTGGCGCGC 1020

QY 1021 TGGGTGTGCGCGACCGCTATCAGGACATAGGCTTGGCTACCGCTGATATTGCTGAAGAGC 1080
Db 1021 TGGGTGTGCGCGACCGCTATCAGGACATAGGCTTGGCTACCGCTGATATTGCTGAAGAGC 1080

QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCGCCGCTCCGATTCGC 1140
Db 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCGCCGCTCCGATTCGC 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTCTGAGTTTAAACAGACCAACG 1200
Db 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTCTGAGTTTAAACAGACCAACG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTTAAAGTACTGGC 1260
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTTAAAGTACTGGC 1260

QY 1261 CGAAGCGCTTGAATAGCGCGGTGCTGCTTCTATATGTTATTTTCCACCATTTG 1320
Db 1261 CGAAGCGCTTGAATAGCGCGGTGCTGCTTCTATATGTTATTTTCCACCATTTG 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTCTCTTCTTACGAGCATTCCT 1380
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTCTCTTCTTACGAGCATTCCT 1380

QY 1381 AGGGGTCTTTCCCTCTGCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGA 1440
Db 1381 AGGGGTCTTTCCCTCTGCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGA 1440

QY 1441 GTTCTCTGGAAGCTTCTTGAAGCAACACGCTCTGTAGCGACCTTTCAGGAGCGG 1500
Db 1441 GTTCTCTGGAAGCTTCTTGAAGCAACACGCTCTGTAGCGACCTTTCAGGAGCGG 1500

QY 1501 AACCCCCACCTGCGGAGCAGGTGCTCTGCGGCGCAAAAGCCACGCTGTATAGATACACCT 1560
Db 1501 AACCCCCACCTGCGGAGCAGGTGCTCTGCGGCGCAAAAGCCACGCTGTATAGATACACCT 1560

QY 1561 GCAAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGAATGTTGGAAGAGTCAAA 1620
Db 1561 GCAAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGAATGTTGGAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCTATTCAACAAGGGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680
Db 1621 TGGCTCTCTCAAGCTATTCAACAAGGGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680

QY 1681 ATGGGATCTGATCTGGGGCCTTCGGTGCACATGCTTTTACATGTTTGTAGTTCGAGGTTAAA 1740
```

[illegible]

Db	2761	CTGCCACC	CGCTAC	GCCTCCGGG	ATCGGTCA	CCGTCGCC	ACATCA	AAACAT	CGAGAG	GTG	2820
Qy	2821	GCTCTG	TCAGCA	CTGGAGAA	ATCCCTTT	TATGCG	AAAGCC	ATCCCA	TCGAGAC	CAATC	2880
Db	2821	GCTCTG	TCAGCA	CTGGAGAA	ATCCCTTT	TATGCG	AAAGCC	ATCCCA	TCGAGAC	CAATC	2880
Qy	2881	AAGGGGG	GAGGAC	CTCATTT	CTGCCA	TTCCA	AGAAAT	TGTAT	GAGCT	CGCCGG	2940
Db	2881	AAGGGGG	GAGGAC	CTCATTT	CTGCCA	TTCCA	AGAAAT	TGTAT	GAGCT	CGCCGG	2940
Qy	2941	AAGCTG	CCGGCC	TCGAGC	TAATGCT	GTAGCA	TATT	ACCGGG	CGCTTG	TATGAT	3000
Db	2941	AAGCTG	CCGGCC	TCGAGC	TAATGCT	GTAGCA	TATT	ACCGGG	CGCTTG	TATGAT	3000
Qy	3001	ATACCA	CTAGCG	GAGAC	GTCAAT	TGTCGT	AGCAAT	TAT	ACCGGG	CGCTTG	3060
Db	3001	ATACCA	CTAGCG	GAGAC	GTCAAT	TGTCGT	AGCAAT	TAT	ACCGGG	CGCTTG	3060
Qy	3061	GGGAT	TTTGAC	TCAGTG	CGATAC	ATATGT	GTAC	CCAGAC	GAGTCG	ACTTCAGC	3120
Db	3061	GGGAT	TTTGAC	TCAGTG	CGATAC	ATATGT	GTAC	CCAGAC	GAGTCG	ACTTCAGC	3120
Qy	3121	CTGAC	CCGAC	CTTCAC	CAAT	TGAGAC	GACAC	CGTG	CCACA	AGCGGT	3180
Db	3121	CTGAC	CCGAC	CTTCAC	CAAT	TGAGAC	GACAC	CGTG	CCACA	AGCGGT	3180
Qy	3181	CAGCGG	CGAGC	AGGACT	GTGTAG	GGG	CAGAT	GGGCAT	TTTAC	AGGTTTG	3240
Db	3181	CAGCGG	CGAGC	AGGACT	GTGTAG	GGG	CAGAT	GGGCAT	TTTAC	AGGTTTG	3240
Qy	3241	GAACGG	CCCTCG	GGGCAT	GTTCGAT	TTCTC	CGTTCT	GTG	CGAGTG	CTATGA	3300
Db	3241	GAACGG	CCCTCG	GGGCAT	GTTCGAT	TTCTC	CGTTCT	GTG	CGAGTG	CTATGA	3300
Qy	3301	GCTTGT	GTACAG	CTCAG	CCCGG	AGAC	CTCAG	TAGT	TGCGG	CTTAC	3360
Db	3301	GCTTGT	GTACAG	CTCAG	CCCGG	AGAC	CTCAG	TAGT	TGCGG	CTTAC	3360
Qy	3361	CCAGGT	TGCGGT	CTGCC	AGGAC	CACTCG	AGT	TTCTGG	GAGAC	GGTCTT	3420
Db	3361	CCAGGT	TGCGGT	CTGCC	AGGAC	CACTCG	AGT	TTCTGG	GAGAC	GGTCTT	3420
Qy	3421	ACCACAT	AGAC	CGCCAT	TTCTT	GTCC	AGACTA	AGCAG	CGAGG	ACAACT	3480
Db	3421	ACCACAT	AGAC	CGCCAT	TTCTT	GTCC	AGACTA	AGCAG	CGAGG	ACAACT	3480
Qy	3481	CTGTG	ATCAT	CAGGCT	ACGGTGT	CGCC	AGGCTC	AGGCT	CCACCT	CCACT	3540
Db	3481	CTGTG	ATCAT	CAGGCT	ACGGTGT	CGCC	AGGCTC	AGGCT	CCACCT	CCACT	3540
Qy	3541	CAAA	TGTGGA	AGTGTCT	ATAC	GGCTAA	AGCTAC	CGT	CGG	CAAC	3600
Db	3541	CAAA	TGTGGA	AGTGTCT	ATAC	GGCTAA	AGCTAC	CGT	CGG	CAAC	3600
Qy	3601	TATAG	CTGG	GAGCC	GTTC	CAAAA	CGAG	GTCTA	CCAC	ACAC	3660
Db	3601	TATAG	CTGG	GAGCC	GTTC	CAAAA	CGAG	GTCTA	CCAC	ACAC	3660
Qy	3661	ATGCA	TGCAT	GTG	CGCTG	ACCTG	GAGG	TGTC	ACG	AGC	3720
Db	3661	ATGCA	TGCAT	GTG	CGCTG	ACCTG	GAGG	TGTC	ACG	AGC	3720
Qy	3721	GTCC	TAC	GAGT	CTG	GGCG	GTATT	GTG	CTG	CAAC	3780
Db	3721	GTCC	TAC	GAGT	CTG	GGCG	GTATT	GTG	CTG	CAAC	3780
Qy	3781	ATCAT	TTTGT	CCGAA	ACCGG	CCAT	CAIT	TTCC	GAC	AGGAA	3840
Db	3781	ATCAT	TTTGT	CCGAA	ACCGG	CCAT	CAIT	TTCC	GAC	AGGAA	3840
Qy	3841	GATGAG	ATGGA	GAGT	GGCT	CTACA	CTCC	CTTT	ATAT	CGAA	3900
Db	3841	GATGAG	ATGGA	GAGT	GGCT	CTACA	CTCC	CTTT	ATAT	CGAA	3900

QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
DB 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
QY 3961 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTCTGGSCGAAGCATATG 4020  
DB 3961 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTCTGGSCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGTTGTCCACTCTGCTGGCAACCC 4080  
DB 4021 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGTTGTCCACTCTGCTGGCAACCC 4080  
QY 4081 GCGATAGCATCACTGATGSCATTCACAGCTCTATCACAGCCCGCTCACACCAACAT 4140  
DB 4081 GCGATAGCATCACTGATGSCATTCACAGCTCTATCACAGCCCGCTCACACCAACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGCGGGGATGGGTGGCGCCCAACTTGTCTCCCAAGCGCT 4200  
DB 4141 ACCCTCTGTTTAAACATCTGCGGGGATGGGTGGCGCCCAACTTGTCTCCCAAGCGCT 4200  
QY 4201 GCTTCTGCTTTCGTAGGCGCGGATCGTGGAGCGCTGTGGCAGCATAGGCGCTTGGG 4260  
DB 4201 GCTTCTGCTTTCGTAGGCGCGGATCGTGGAGCGCTGTGGCAGCATAGGCGCTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGSGTGGCAGGCGCGCTCGTGCC 4320  
DB 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGSGTGGCAGGCGCGCTCGTGCC 4320  
QY 4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACGAGGACCTGGTTAACTACTCCCTGCT 4380  
DB 4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACGAGGACCTGGTTAACTACTCCCTGCT 4380  
QY 4381 ATCTCTCCCTGGCGCCTAGTCTCGGGGTGCTGGCAGCGATACTCGTGGCAC 4440  
DB 4381 ATCTCTCCCTGGCGCCTAGTCTCGGGGTGCTGGCAGCGATACTCGTGGCAC 4440  
QY 4441 GTGGGCCACAGGAGGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
DB 4441 GTGGGCCACAGGAGGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
QY 4501 GGTAAACACGCTCTCCCAACGCTATGTGCTGAGAGCGAGCTGCAAGCAGTGTCACT 4560  
DB 4501 GGTAAACACGCTCTCCCAACGCTATGTGCTGAGAGCGAGCTGCAAGCAGTGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGAGTCAAC 4620  
DB 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGAGTCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
DB 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCCTGSCCGGATTTGCCGGGA 4740  
DB 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCCTGSCCGGATTTGCCGGGA 4740  
QY 4741 GTCCCTCTTCTCATGTCAAGTGGGTACAAAGGAGTCTGGCGGGGAGCGGATCATG 4800  
DB 4741 GTCCCTCTTCTCATGTCAAGTGGGTACAAAGGAGTCTGGCGGGGAGCGGATCATG 4800  
QY 4801 CAAACCACTGCCATGTGGAGCAGATCACCGGACATGTGAAACCGTTCCATGAGG 4860  
DB 4801 CAAACCACTGCCATGTGGAGCAGATCACCGGACATGTGAAACCGTTCCATGAGG 4860  
QY 4861 ATCTGGGGGCTAGACCTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACGCGTAC 4920  
DB 4861 ATCTGGGGGCTAGACCTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACGCGTAC 4920  
QY 4921 ACCAGGCGCCCTGACGCGCCCTCCCGGCGCAAAATTTCTAGGGGCTGTGGCGGGTG 4980  
DB 4921 ACCAGGCGCCCTGACGCGCCCTCCCGGCGCAAAATTTCTAGGGGCTGTGGCGGGTG 4980

QY 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTAGCTAGCGGCGCATG 5040  
DB 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTAGCTAGCGGCGCATG 5040  
QY 5041 ACCACTGACAAAGTAAAGTGGCCGTTGTCAGGTTCCGGCCCCCGAATTTCTTACACAAGTG 5100  
DB 5041 ACCACTGACAAAGTAAAGTGGCCGTTGTCAGGTTCCGGCCCCCGAATTTCTTACACAAGTG 5100  
QY 5101 GATGGGTGCGGTTGTCACACAGTACCGCTCCAGGTCGCAACCCCTCTTACGGGAGAGGTC 5160  
DB 5101 GATGGGTGCGGTTGTCACACAGTACCGCTCCAGGTCGCAACCCCTCTTACGGGAGAGGTC 5160  
QY 5161 ACATTTCTGCTGCGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGAGCGCCGAA 5220  
DB 5161 ACATTTCTGCTGCGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGAGCGCCGAA 5220  
QY 5221 CCGGACGTAGCAGTGTCTCACTTCCATGTCTACCGACCCCTCCCATTTACGGCGGAGAGC 5280  
DB 5221 CCGGACGTAGCAGTGTCTCACTTCCATGTCTACCGACCCCTCCCATTTACGGCGGAGAGC 5280  
QY 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
DB 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
QY 5341 CTGCTGCGCCTTCTTGAAGGCAACATGCACTACCCGCTCATGACTCCCCCGACGCTGAC 5400  
DB 5341 CTGCTGCGCCTTCTTGAAGGCAACATGCACTACCCGCTCATGACTCCCCCGACGCTGAC 5400  
QY 5401 CTCATCGAGGCAACCTCTCTGTGGCGGAGGAGATGGCGGGGAAATCACCCCGCTGGAG 5460  
DB 5401 CTCATCGAGGCAACCTCTCTGTGGCGGAGGAGATGGCGGGGAAATCACCCCGCTGGAG 5460  
QY 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTGGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
DB 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTGGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGACGAGT 5580  
DB 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGACGAGT 5580  
QY 5581 CCCATATGGGCACGCCCGGATTACAAACCTCCACTGTTTAGAGTCTCTGAAGGACCCCGAC 5640  
DB 5581 CCCATATGGGCACGCCCGGATTACAAACCTCCACTGTTTAGAGTCTCTGAAGGACCCCGAC 5640  
QY 5641 TAGTCCCTCAGTGGTACACGGGTGTCATTCGCGCTCCAAAGCCCTCTCCGATACCA 5700  
DB 5641 TAGTCCCTCAGTGGTACACGGGTGTCATTCGCGCTCCAAAGCCCTCTCCGATACCA 5700  
QY 5701 CCTCCAGGAGAGAGAGCGGTTGCTGTCAGAACTCTACCGTGTCTTCTGCTTGGCG 5760  
DB 5701 CCTCCAGGAGAGAGAGCGGTTGCTGTCAGAACTCTACCGTGTCTTCTGCTTGGCG 5760  
QY 5761 GAGTCCGCCACAAAGACCTTCGGCAGTTCGAAATCGTGGCGGCTCGACAGCGGCACGCA 5820  
DB 5761 GAGTCCGCCACAAAGACCTTCGGCAGTTCGAAATCGTGGCGGCTCGACAGCGGCACGCA 5820  
QY 5821 ACGGCTCTCTGACAGCCCTCCGACGCGGACGCGGATCCGAGCTTGAGTCTGATC 5880  
DB 5821 ACGGCTCTCTGACAGCCCTCCGACGCGGACGCGGATCCGAGCTTGAGTCTGATC 5880  
QY 5881 TCCTTCCATGCCCGCTTGGCGGAGCGCGGGGATCCCGATCTCAGCGCGGCTTGG 5940  
DB 5881 TCCTTCCATGCCCGCTTGGCGGAGCGCGGGGATCCCGATCTCAGCGCGGCTTGG 5940  
QY 5941 TCTACCGTAGCGGAGGCTTAGTGAAGACGTGCTGTGCTGCTGATGTCTTACATGG 6000  
DB 5941 TCTACCGTAGCGGAGGCTTAGTGAAGACGTGCTGTGCTGCTGATGTCTTACATGG 6000  
QY 6001 ACAGCGCCCTGATCACGCCATGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
DB 6001 ACAGCGCCCTGATCACGCCATGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAACTCTCGCAGCGCAAGC 6120





FEATURES		Location/Qualifiers	
source		1..8001	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match 99.7%; Score 7965.4; DB 6; Length 8001;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 7988; Conservative 0; Mismatches 1; Indels 12; Gaps 1;			
QY	1	GCACGCCCGGATTGGGGCGGACATCCACCATAGATCACTCCCTCTGTGAGGAATCTACTG	60
DB	1	GCACGCCCGGATTGGGGCGGACATCCACCATAGATCACTCCCTCTGTGAGGAATCTACTG	60
QY	61	TCCTCAGCAGAAAGCGTCTAGCCATGGGTAGTATGATGATGCTGCTGAGCCTCCAGGAC	120
DB	61	TCCTCAGCAGAAAGCGTCTAGCCATGGGTAGTATGATGATGCTGCTGAGCCTCCAGGAC	120
QY	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG	180
DB	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG	180
QY	181	GACGACCGGTCCTTCTTGTGATCAACCGGCTCAATGCTGAGATTTGGGGTGCCTCC	240
DB	181	GACGACCGGTCCTTCTTGTGATCAACCGGCTCAATGCTGAGATTTGGGGTGCCTCC	240
QY	241	CGGAGCTGCTAGCGGAGTAGTGGTTCGCGAAGGCTTGTGGTACTGCTGATAGG	300
DB	241	CGGAGCTGCTAGCGGAGTAGTGGTTCGCGAAGGCTTGTGGTACTGCTGATAGG	300
QY	301	GTGCTTGCAGTGCCCGGAGGCTCTGTAGACCGGTGACCATGAGCAGCAATCTTAAC	360
DB	301	GTGCTTGCAGTGCCCGGAGGCTCTGTAGACCGGTGACCATGAGCAGCAATCTTAAC	360
QY	361	CTCAAGAAACCAAA-----GGCGGCGCATGATTGAAACAAGATGATTGC	408
DB	361	CTCAAGAAACCAAA-----GGCGGCGCATGATTGAAACAAGATGATTGC	420
QY	409	ACGAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTGGGTATGACTGCGGCACACAGA	468
DB	421	ACGAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTGGGTATGACTGCGGCACACAGA	480
QY	469	CAATCGGCTGCTGATGCCCGGCTGTCAGCGGCGGCGCGCGGCTCTT	528
DB	481	CAATCGGCTGCTGATGCCCGGCTTCCGGCTGTCAGCGGCGGCGCGGCTCTT	540
QY	529	TTGTCAAGACGACCTGTCCGGTCCCTGATGAATGAACTGACGAGCAGCGGGCTAT	588
DB	541	TTGTCAAGACGACCTGTCCGGTCCCTGATGAATGAACTGACGAGCAGCGGGCTAT	600
QY	589	CGTGGCTGGCCACGACGGGCGTTCCTTCCGCGAGCTGTCTGACGCTTCTCACTGAAGCGG	648
DB	601	CGTGGCTGGCCACGACGGGCGTTCCTTCCGCGAGCTGTCTGACGCTTCTCACTGAAGCGG	660
QY	649	GAAAGGACTGCTGCTATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTCACTCACTTG	708
DB	661	GAAAGGACTGCTGCTATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTCACTCACTTG	720
QY	709	CTCTCTCGGAGAAAGTATCCATCATGCTGATGATCAATGCGGCGGCTGATACGCTGATC	768
DB	721	CTCTCTCGGAGAAAGTATCCATCATGCTGATGATCAATGCGGCGGCTGATACGCTGATC	780
QY	769	CGGTACTCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTACTCGGA	828
DB	781	CGGTACTCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTACTCGGA	840
QY	829	TGGAAGCGGCTTGTGATCAGATGATCTGGAAGAGCATCAGGGGCTCGCGCAG	888
DB	841	TGGAAGCGGCTTGTGATCAGATGATCTGGAAGAGCATCAGGGGCTCGCGCAG	900
QY	889	CCGAACCTTTCGAGGCTCAAGGCGCATGCGCGGCGAGGATCTCGTCTGACCC	948
DB	901	CCGAACCTTTCGAGGCTCAAGGCGCATGCGCGGCGAGGATCTCGTCTGACCC	960
QY	949	ATGGCGATGCCCTGCTTGGCGAATATATGCTGGAATAATGCGCGCTTCTTGGATTGATCG	1008
DB	961	ATGGCGATGCCCTGCTTGGCGAATATATGCTGGAATAATGCGCGCTTCTTGGATTGATCG	1020
QY	1009	ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATACGCTTGGCTACCGGTGATA	1068
DB	1021	ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATACGCTTGGCTACCGGTGATA	1080
QY	1069	TTGTGAAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTGCTTACGGTATCGCGG	1128
DB	1081	TTGTGAAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTGCTTACGGTATCGCGG	1140
QY	1129	CTCCCGATTCCGACGCGATCGCTTCTATCGCTTCTTGAACGATCTCTGAGTTTAAA	1188
DB	1141	CTCCCGATTCCGACGCGATCGCTTCTATCGCTTCTTGAACGATCTCTGAGTTTAAA	1200
QY	1189	CAGACCAACACGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTCT	1248
DB	1201	CAGACCAACACGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTCT	1260
QY	1249	AACGTTACTGGCGGAGCGGCTTGAATTAAGGCGGCTGTGGCTTGTCTATATGTTATTT	1308
DB	1261	AACGTTACTGGCGGAGCGGCTTGAATTAAGGCGGCTGTGGCTTGTCTATATGTTATTT	1320
QY	1309	TCCACCATATTGCCGCTTCTTTGGCAATGCTGAGGCGCGGAAACCTGGCGCTGTCTCTTG	1368
DB	1321	TCCACCATATTGCCGCTTCTTTGGCAATGCTGAGGCGCGGAAACCTGGCGCTGTCTCTTG	1380
QY	1369	ACGAGCATTCCTAGGGGTCTTTCCCTCTCGCAAGGAATGCAAGTCTGTGTAATGTC	1428
DB	1381	ACGAGCATTCCTAGGGGTCTTTCCCTCTCGCAAGGAATGCAAGTCTGTGTAATGTC	1440
QY	1429	GTGAAGAGAGCAGTTCTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCTT	1488
DB	1441	GTGAAGAGAGCAGTTCTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCTT	1500
QY	1489	TGCGAGGACGGGAAACCCCGACCTGGGACAGGTGCTCTGGCGCAAAAGCCAGCTGTA	1548
DB	1501	TGCGAGGACGGGAAACCCCGACCTGGGACAGGTGCTCTGGCGCAAAAGCCAGCTGTA	1560
QY	1549	TAAGATACCTCTCAAGAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGATAGTTGTG	1608
DB	1561	TAAGATACCTCTCAAGAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGATAGTTGTG	1620
QY	1609	GAAAGAGTCAAAATGGCTCTCTCAAGGTAATTCAAAGAGGCTGAAGGATGCCAGAG	1668
DB	1621	GAAAGAGTCAAAATGGCTCTCTCAAGGTAATTCAAAGAGGCTGAAGGATGCCAGAG	1680
QY	1669	GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTTTTAG	1728
DB	1681	GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTTTTAG	1740
QY	1729	TCGAGGTTAAAAACGCTTAGGCCCCCGAACCAACGAGGGAAGTGGTTTTCTTTGAAAAA	1788
DB	1741	TCGAGGTTAAAAACGCTTAGGCCCCCGAACCAACGAGGGAAGTGGTTTTCTTTGAAAAA	1800
QY	1789	CACGATTAATCCATGGCGCTTATTCAGGCTTATCTCCCAACAGACGCGAGGCTTCTTGGC	1848
DB	1801	CACGATTAATCCATGGCGCTTATTCAGGCTTATTCAGGCTTATCTCCCAACAGACGCGAGGCTTCTTGGC	1860
QY	1849	TGCATCATCTAGCTCAGAGCGGCGGACAGGAAACAGGTTCGAGGGGAGGTCCAGTG	1908
DB	1861	TGCATCATCTAGCTCAGAGCGGCGGACAGGAAACAGGTTCGAGGGGAGGTCCAGTG	1920
QY	1909	GTCTCCACCGCAACACAATCTTCTGGCGACCTTCGCTCAATGCGCTGTGTGAGCTGTC	1968
DB	1921	GTCTCCACCGCAACACAATCTTCTGGCGACCTTCGCTCAATGCGCTGTGTGAGCTGTC	1980
QY	1969	TATCATGGTCCGCTCAAGAGACCTTTCGCGGCGCAAGGCGCCCAATCACCACCAATGTAC	2028
DB	1981	TATCATGGTCCGCTCAAGAGACCTTTCGCGGCGCAAGGCGCCCAATCACCACCAATGTAC	2040



2029 QY ACCAATGTGGACAGGAACCTCTCGGTGGGAAAGCGCCCCCGGGGGCGGTTCCTTGACA 2088  
2041 Db ACCAATGTGGACAGGAACCTCTCGGTGGGAAAGCGCCCCCGGGGGCGGTTCCTTGACA 2100  
2089 QY CCAATGCACCTCGGGAGCTCGGACCTTTACTTTGGTCAAGGAGCATGSCGATGTCATTCCG 2148  
2101 Db CCAATGCACCTCGGGAGCTCGGACCTTTACTTTGGTCAAGGAGCATGCCGATGTCAITCCG 2160  
2149 QY GTGCGCGCGGGGGGACAGCAGGGGAGCCTACTCTCCCCCAGGCGCGCTCTCCTACTTG 2208  
2161 Db GTGCGCGCGGGGGACAGCAGGGGAGCCTACTCTCCCCCAGGCGCGCTCTCCTACTTG 2220  
2209 QY AAGGGCTCTTTCGGGGGCTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCACTTTCCG 2268  
2221 Db AAGGGCTCTTTCGGGGGCTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCACTTTCCG 2280  
2269 QY GCTGCCGTGTGCACCGGAGGGGTGGAAAGGGGTGGACTTTGTACCGCTCGAGTCTATG 2328  
2281 Db GCTGCCGTGTGCACCGGAGGGGTGGAAAGGGGTGGACTTTGTACCGCTCGAGTCTATG 2340  
2329 QY GAAACCACTATGCGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGCGCGTACCGCAG 2388  
2341 Db GAAACCACTATGCGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGCGCGTACCGCAG 2400  
2389 QY ACATTCAGGTGCCCATCTACAGCGCCCTACTGTGTAGCGGCAAGACATAAGTGGCG 2448  
2401 Db ACATTCAGGTGCCCATCTACAGCGCCCTACTGTGTAGCGGCAAGACATAAGTGGCG 2460  
2449 QY GCTGCCGTGTGCACCGGAGGTATAGGTGCTTGTCTGAAACCGCTCCGTCGCGCGCAC 2508  
2461 Db GCTGCCGTGTGCACCGGAGGTATAGGTGCTTGTCTGAAACCGCTCCGTCGCGCGCAC 2520  
2509 QY CTAGTTCCTGGGCGGTATGTCTAAGGCACATGATCGACCTAATCAATCAAGACCGGG 2568  
2521 Db CTAGTTCCTGGGCGGTATGTCTAAGGCACATGATCGACCTAATCAATCAAGACCGGG 2580  
2569 QY GTAAGGACCATACACAGGGTGCCTCATCAGTACTCAGCTATGCGCAAGTTCCTGSC 2628  
2581 Db GTAAGGACCATACACAGGGTGCCTCATCAGTACTCAGCTATGCGCAAGTTCCTGSC 2640  
2629 QY GAGGTGTGTCTCGGGCGCTATGACATCATATATGTGATGAGTGCCTCACT 2688  
2641 Db GAGGTGTGTCTCGGGCGCTATGACATCATATATGTGATGAGTGCCTCACT 2700  
2689 QY GACTCGACCATATCTCTGGGATCGGCACAGTCTCGAACCGGAGCGGTGGAGCG 2748  
2701 Db GACTCGACCATATCTCTGGGATCGGCACAGTCTCGAACCGGAGCGGTGGAGCG 2760  
2749 QY CGACTGTGTCTCGGCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAAC 2808  
2761 Db CGACTGTGTCTCGGCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAAC 2820  
2809 QY ATCGAGGAGTGTCTGTCCAGCAGTGGAGAAATCCCTTTATGCAAGGCAATCCCT 2868  
2821 Db ATCGAGGAGTGTCTGTCCAGCAGTGGAGAAATCCCTTTATGCAAGGCAATCCCT 2880  
2869 QY ATCGAGACCATCAAGGGGGGAGGCACCTCATTTCTGCCATTCCAAAGAGAAATGTAT 2928  
2881 Db ATCGAGACCATCAAGGGGGGAGGCACCTCATTTCTGCCATTCCAAAGAGAAATGTAT 2940  
2929 QY GAGTCCCGGGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGCCT 2988  
2941 Db GAGTCCCGGGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGCCT 3000  
2989 QY GATGTATCCGTATACCACTAGCGGAGAGCTCATTTGTGTAGCAACCGGCGCTCTAATG 3048  
3001 Db GATGTATCCGTATACCACTAGCGGAGAGCTCATTTGTGTAGCAACCGGCGCTCTAATG 3060  
3049 QY ACGGGCTTTACCGGCGATTTTCGACTCAGTGTATCGACTGCAATACATGTGTACCCAGACA 3108  
3061 Db ACGGGCTTTACCGGCGATTTTCGACTCAGTGTATCGACTGCAATACATGTGTACCCAGACA 3120  
3109 QY GTCGACTTTCAGCTCGGACCCGACCTTCACCATTTGAGACGAGCCGTGCGCAAGACGG 3168

3121 Db GTGCACTTCAGCTCGAACCCGACCTTCACATTTGAGACGACGACCTGCGCAACAAGCGG 3180  
3169 QY GTGTCACTCGTCCAGCGGAGGAGGACTGTGTAGGGGCAAGATGGGCAATTTACAGTTT 3228  
3181 Db GTGTCACTCGTCCAGCGGAGGAGGACTGTGTAGGGGCAAGATGGGCAATTTACAGTTT 3240  
3229 QY GTGACTCCAGGAGAACGGCCCTCGGGCATGTTGATTCCTCGGTCTGTGCGAGTCTAT 3288  
3241 Db GTGACTCCAGGAGAACGGCCCTCGGGCATGTTGATTCCTCGGTCTGTGCGAGTCTAT 3300  
3289 QY GACGGGGCTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGGGCT 3348  
3301 Db GACGGGGCTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGGGCT 3360  
3349 QY TACTTAAACACACCAAGGTTGCCCTGTGCCAGGACCATCTGGAGTTCTGGGAGAGCTC 3408  
3361 Db TACTTAAACACACCAAGGTTGCCCTGTGCCAGGACCATCTGGAGTTCTGGGAGAGCTC 3420  
3409 QY TTTTACGGCTCACCCACATAGACGCCCATTTTCTTGTCCAGACTAAGCAGGAGAGAC 3468  
3421 Db TTTTACGGCTCACCCACATAGACGCCCATTTTCTTGTCCAGACTAAGCAGGAGAGAC 3480  
3469 QY AACTTCCCCTACTGCTAGCATACAGGCTACGGTGTGCCCGCAGGGCTCAGGCTCCACCT 3528  
3481 Db AACTTCCCCTACTGCTAGCATACAGGCTACGGTGTGCCCGCAGGGCTCAGGCTCCACCT 3540  
3529 QY CCATCTGGGACCAATGTGGAAGTGTCTATACGCTTAAAGCTTAAGCTGACGCGGCA 3588  
3541 Db CCATCTGGGACCAATGTGGAAGTGTCTATACGCTTAAAGCTTAAGCTGACGCGGCA 3600  
3589 QY AGCCCTCTCTATAGGCTGGGAGCCGTTCAAAAGAGGTTACTACACACACCCCAT 3648  
3601 Db AGCCCTCTCTATAGGCTGGGAGCCGTTCAAAAGAGGTTACTACACACACCCCAT 3660  
3649 QY ACCAAATACATCATGTCATGCTCGGTGACCTCGAGGTCGTACGAGCACTCGGCTG 3708  
3661 Db ACCAAATACATCATGTCATGCTCGGTGACCTCGAGGTCGTACGAGCACTCGGCTG 3720  
3709 QY CTGTTAGCGGAGTCTAGCAGCTGTGGCGGATGTTGCTGTGACAAAGGAGCGGTGTC 3768  
3721 Db CTGTTAGCGGAGTCTAGCAGCTGTGGCGGATGTTGCTGTGACAAAGGAGCGGTGTC 3780  
3769 QY ATTGTGGGAGGATCATCTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCCCT 3828  
3781 Db ATTGTGGGAGGATCATCTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCCCT 3840  
3829 QY TACCGGAGTTCGATGAGATGGAAGTGGCGCTCACCTCCCTTACATCGAACAGGGA 3888  
3841 Db TACCGGAGTTCGATGAGATGGAAGTGGCGCTCACCTCCCTTACATCGAACAGGGA 3900  
3889 QY ATGAGCTCGCGAAACAAATTCAAACAGAGGCAATCGGTTGTGCAAAAGCAGCACCAG 3948  
3901 Db ATGAGCTCGCGAAACAAATTCAAACAGAGGCAATCGGTTGTGCAAAAGCAGCACCAG 3960  
3949 QY CAAGCGGAGCTGCTGCTCCGTTGGTGAATCCAAAGTGGGAGCCCTCGAGGCTTCTG 4008  
3961 Db CAAGCGGAGCTGCTGCTCCGTTGGTGAATCCAAAGTGGGAGCCCTCGAGGCTTCTG 4020  
4009 QY GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4068  
4021 Db GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4080  
4069 QY CTTGGCAACCCCGCATAGCATCACTGATGGCAATTCACAGCTCTATACACAGCGGCTC 4128  
4081 Db CTTGGCAACCCCGCATAGCATCACTGATGGCAATTCACAGCTCTATACACAGCGGCTC 4140  
4129 QY ACCACCAACATACCTCTCTGTTTAAATCCTCGGGGATGGGTGCGCGCCCACTGCT 4188  
4141 Db ACCACCAACATACCTCTCTGTTTAAATCCTCGGGGATGGGTGCGCGCCCACTGCT 4200  
4189 QY CTTCCAGCGCTCTCTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTGTGTCAGC 4248

Db	4201	CCTCCGACGCTGCTTCTGCTTTTCGTAGCGCCGGCAGTCGCTGAGACGGCTGTGTGGCAGC	4261
QY	4249	ATAGCCCTTGGAAAGTGCTTTGTGGATATTTTGGCAGCTTATGAGACAGGGGTGGCAGGC	4308
Db	4261	ATAGCCCTTGGAAAGTGCTTTGTGGATATTTTGGCAGCTTATGAGACAGGGGTGGCAGGC	4320
QY	4309	GGCTCTGTGGCTTTAAGTTCATGAGCGGCGAGATGCCCTCCACCGAGGACCTTGGTTAAC	4368
Db	4321	GGCTCTGTGGCTTTAAGTTCATGAGCGGCGAGATGCCCTCCACCGAGGACCTTGGTTAAC	4380
QY	4369	CTACTCCCTGCTATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCTGTGTGGCAGCGATA	4428
Db	4381	CTACTCCCTGCTATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCTGTGTGGCAGCGATA	4440
QY	4429	CTGGCTCGGCACGTGGGCCCCAGGGAGGGGGCTCTGCAGTGGATGAACCCGGCTGATAGCG	4488
Db	4441	CTGGCTCGGCACGTGGGCCCCAGGGAGGGGGCTCTGCAGTGGATGAACCCGGCTGATAGCG	4500
QY	4489	TTCCGTTTCGGGGGTAAACCAAGTCTCCCCACGCACTATGTGCTTGAGACGCGCTGCA	4548
Db	4501	TTCCGTTTCGGGGGTAAACCAAGTCTCCCCACGCACTATGTGCTTGAGACGCGCTGCA	4560
QY	4549	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC	4608
Db	4561	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC	4620
QY	4609	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCTGGGCTTAAGAGATGTTGG	4668
Db	4621	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCTGGGCTTAAGAGATGTTGG	4680
QY	4669	GATTGGATATGACCGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTCCCG	4728
Db	4681	GATTGGATATGACCGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTCCCG	4740
QY	4729	CGATTGCGGGAGTCCCTTCTCTCATGTCAAGTGGGTAAAGGGAGTCTGGCGGGC	4788
Db	4741	CGATTGCGGGAGTCCCTTCTCTCATGTCAAGTGGGTAAAGGGAGTCTGGCGGGC	4800
QY	4789	GAGGCATCATGCAAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTGAAAAAC	4848
Db	4801	GAGGCATCATGCAAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTGAAAAAC	4860
QY	4849	GGTTCCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGAACTTCCCG	4908
Db	4861	GGTTCCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGAACTTCCCG	4920
QY	4909	ATTAAACGCTACACCGGGCCCTGCAGCCCTCCCGGGCGCAATTTATTTCTAGGGCG	4968
Db	4921	ATTAAACGCTACACCGGGCCCTGCAGCCCTCCCGGGCGCAATTTATTTCTAGGGCG	4980
QY	4969	CTGTGGCGGTGCTGTGTAGGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC	5028
Db	4981	CTGTGGCGGTGCTGTGTAGGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC	5040
QY	5029	GTACGGGCATGACCACTGACAAAGTGTCCCGTGTCCAGTTCCGGCCCCCGAATTC	5088
Db	5041	GTACGGGCATGACCACTGACAAAGTGTCCCGTGTCCAGTTCCGGCCCCCGAATTC	5100
QY	5089	TTCAACAGATGATGGGGTGGCGTTGCAAGTACGCTCCAGCGTGCAAAACCCCTCTTA	5148
Db	5101	TTCAACAGATGATGGGGTGGCGTTGCAAGTACGCTCCAGCGTGCAAAACCCCTCTTA	5160
QY	5149	CGGGAGAGGTACATTCCTTGTGCGGGCTCAATCAATACCTGTTGGGTACAGCTCCA	5208
Db	5161	CGGGAGAGGTACATTCCTTGTGCGGGCTCAATCAATACCTGTTGGGTACAGCTCCA	5220
QY	5209	TGGAGCCGGAACCGGACGTAGCTCTCACTTCCATGCTCACCGACCCCTCCACATT	5268
Db	5221	TGGAGCCGGAACCGGACGTAGCTCTCACTTCCATGCTCACCGACCCCTCCACATT	5280
QY	5269	ACGGCGGACAGCGCTTAAGCTTAGCTTGGCCAGGGGATCTCCCCCTCTTTGGCCAGCTCA	5328
Db	5281	ACGGCGGACAGCGCTTAAGCTTAGCTTGGCCAGGGGATCTCCCCCTCTTTGGCCAGCTCA	5340

QY	5329	TCAGCTACGAGCTGTCTGCGCCTTCCTTGAAGGCAACATGCACTACCCGTCATGACTCC	5388
DB	5341	TCAGCTACCGAGCTGTCTGCGCCTTCCTTGAAGGCAACATGCACTACCCGTCATGACTCC	5400
QY	5389	CCGAGCGCTGACCTCATCGAGGCCAACTCTCTGTGGCGGAGGAGATGGCGGGAAATC	5448
DB	5401	CCGAGCGCTGACCTCATCGAGGCCAACTCTCTGTGGCGGAGGAGATGGCGGGAAATC	5460
QY	5449	ACCCGCGTGGAGTCAGAAAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCG	5508
DB	5461	ACCCGCGTGGAGTCAGAAAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCG	5520
QY	5509	GAGGAGGATGAGGGAGATACCGTTCCGGCGGAGATCCTGCGGAGGTCAGGAAATTC	5568
DB	5521	GAGGAGGATGAGGGAGATACCGTTCCGGCGGAGATCCTGCGGAGGTCAGGAAATTC	5580
QY	5569	CCTCGAGCGATGCCCATATGGCAGCGCCCGGATTAAACCCCTCCACTGTTAGAGTCCCTGG	5628
DB	5581	CCTCGAGCGATGCCCATATGGCAGCGCCCGGATTAAACCCCTCCACTGTTAGAGTCCCTGG	5640
QY	5629	AAGACCCGACTAGTCCCTCAGTGGTACACGGGTGTCATTGCGGCTGTCGAAGGCC	5688
DB	5641	AAGACCCGACTAGTCCCTCAGTGGTACACGGGTGTCATTGCGGCTGTCGAAGGCC	5700
QY	5689	CCTCCGATACCACTCCACGAGGAAGAGACGGTTGTCTCTGTCAGAACTACCGGTGCT	5748
DB	5701	CCTCCGATACCACTCCACGAGGAAGAGACGGTTGTCTCTGTCAGAACTACCGGTGCT	5760
QY	5749	TCTGCTTGGCGAGCTGCGCAAAAGACTTCGGCAGCTCGAAATCGTCGGCCGCTCGAC	5808
DB	5761	TCTGCTTGGCGAGCTGCGCAAAAGACTTCGGCAGCTCGAAATCGTCGGCCGCTCGAC	5820
QY	5809	AGCGGACGCGCAAGGCGCTCTCTGACCAAGCCCTCCGACGACGGCGACGGGATCCGAC	5868
DB	5821	AGCGGACGCGCAAGGCGCTCTCTGACCAAGCCCTCCGACGACGGCGACGGGATCCGAC	5880
QY	5869	GTTGAGTCGTACTCCTCCATGCCCCCTTTGAGGGGAGCGGGGGATCCCGATCTCAGC	5928
DB	5881	GTTGAGTCGTACTCCTCCATGCCCCCTTTGAGGGGAGCGGGGGATCCCGATCTCAGC	5940
QY	5929	GACGGCTTTGGTCTACCGTAAGCGAGGAGGCTAGTGAAGACGTGCTGTCTGTCTGATG	5988
DB	5941	GACGGCTTTGGTCTACCGTAAGCGAGGAGGCTAGTGAAGACGTGCTGTCTGTCTGATG	6000
QY	5989	TCCTACATATGACAGGCGCCTGATCAGCCCATGCGCTGCGGAGGAACCAAGCTGCC	6048
DB	6001	TCCTACATATGACAGGCGCCTGATCAGCCCATGCGCTGCGGAGGAACCAAGCTGCC	6060
QY	6049	ATCAATGCACTGAGCAACTCTTTGTCTCGTCAACCAACTTGGTCTATGCTACAACTCT	6108
DB	6061	ATCAATGCACTGAGCAACTCTTTGTCTCGTCAACCAACTTGGTCTATGCTACAACTCT	6120
QY	6109	CGCAGGCAAGCTCGGGCAGAGAGGTCACCTTTGACAGACTCCAGTCTCTGAGACAC	6168
DB	6121	CGCAGGCAAGCTCGGGCAGAGAGGTCACCTTTGACAGACTCCAGTCTCTGAGACAC	6180
QY	6169	CACTACGGGACGTGCTCAAGGAGATGAAGGGGAAGGGCGTCCACAGTTAAGGCTAAAATT	6228
DB	6181	CACTACGGGACGTGCTCAAGGAGATGAAGGGGAAGGGCGTCCACAGTTAAGGCTAAAATT	6240
QY	6229	CTATCCGTGGAGAGCGCTGAAGCTGACGCCGCCACATTCCGCCAGATCTAAATTGGC	6288
DB	6241	CTATCCGTGGAGAGCGCTGAAGCTGACGCCGCCACATTCCGCCAGATCTAAATTGGC	6300
QY	6289	TATGGGCAAGGACGCTCCGGAACCTATCCAGCAAGGCGGTTAAACCATCCGCTCCGCTG	6348
DB	6301	TATGGGCAAGGACGCTCCGGAACCTATCCAGCAAGGCGGTTAAACCATCCGCTCCGCTG	6360
QY	6349	TGGAAGGACTTGCTGGAAGACACTGAGACCAAAATTGACACCAATCATGGCAAAAAAT	6408
DB	6361	TGGAAGGACTTGCTGGAAGACACTGAGACCAAAATTGACACCAATCATGGCAAAAAAT	6420

```
QY 6409 GAGGTTTCTCGGTCCTCAACCGAGAGAGGGGGGCGCCAGCCAGCTCGGCTTATCGTATTC 6468
Db 6421 GAGGTTTCTCGGTCCTCAACCGAGAGAGGGGGGCGCCAGCCAGCTCGGCTTATCGTATTC 6480
QY 6469 CCAGATTTGGGGGTTGCGTGTGCGCAGAAATGGCCCTTACGATGTGCTCTCCACCCCTC 6528
Db 6481 CCAGATTTGGGGGTTGCGTGTGCGCAGAAATGGCCCTTACGATGTGCTCTCCACCCCTC 6540
QY 6529 CCTCAGGCGCGTGATGGGCTCTTATACCGAATCCAAATCTCTCTCGACAGCGGTCGAG 6588
Db 6541 CCTCAGGCGCGTGATGGGCTCTTATACCGAATCCAAATCTCTCTCGACAGCGGTCGAG 6600
QY 6589 TTCCTGTGTAATGCTGTGAAACCGAGAAATGCCCTATAGGCTTCGCATATGACACCCGC 6648
Db 6601 TTCCTGTGTAATGCTGTGAAACCGAGAAATGCCCTATAGGCTTCGCATATGACACCCGC 6660
QY 6649 TGTTTTGACTCAACGGTCACTCAGAAATGATCCCGTGTGAGGAGTCAATCTACCAATGT 6708
Db 6661 TGTTTTGACTCAACGGTCACTCAGAAATGATCCCGTGTGAGGAGTCAATCTACCAATGT 6720
QY 6709 TGTGACTTGGCCCCGGAAGCCAGACAGCCATAAGTCCGCTCAAGAGCGGCTTTACATC 6768
Db 6721 TGTGACTTGGCCCCGGAAGCCAGACAGCCATAAGTCCGCTCAAGAGCGGCTTTACATC 6780
QY 6769 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6828
Db 6781 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6840
QY 6829 GGTGTACTCAGCAGCAGCTGCGGTATACCTCACATGTTACTTTGAAGCCGCTCGCGGC 6888
Db 6841 GGTGTACTCAGCAGCAGCTGCGGTATACCTCACATGTTACTTTGAAGCCGCTCGCGGC 6900
QY 6889 TGTGAGCTCGGAAGCTCCAGAGCTCAGCATGCTCGTATCGGAGACGACCTTGTGTT 6948
Db 6901 TGTGAGCTCGGAAGCTCCAGAGCTCAGCATGCTCGTATCGGAGACGACCTTGTGTT 6960
QY 6949 ATCTGTGAAGCGCGGGGACCCNAGAGGACGAGCGCTACGGGCTTACGGAGGCT 7008
Db 6961 ATCTGTGAAGCGCGGGGACCCNAGAGGACGAGCGCTACGGGCTTACGGAGGCT 7020
QY 7009 ATGACTAGATATCTGCCCCCTCGGGGACCGCCCAACAGAAATACGACTTGGAGTTG 7068
Db 7021 ATGACTAGATATCTGCCCCCTCGGGGACCGCCCAACAGAAATACGACTTGGAGTTG 7080
QY 7069 ATAACATCATGCTCTCCATGTTGTCAGTCGCGCAGCATGCTCGGCAAAAGGGGTGAC 7128
Db 7081 ATAACATCATGCTCTCCATGTTGTCAGTCGCGCAGCATGCTCGGCAAAAGGGGTGAC 7140
QY 7129 TATCTCACCGGTGACCCGACCCCTTGGCGGGCTGCGGGCTGCGGGAGAGCATAGACAC 7188
Db 7141 TATCTCACCGGTGACCCGACCCCTTGGCGGGCTGCGGGCTGCGGGAGAGCATAGACAC 7200
QY 7189 ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGATCGGCCACCTTGTGGCAAGG 7248
Db 7201 ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGATCGGCCACCTTGTGGCAAGG 7260
QY 7249 ATGATCCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACTTGAAGAAGC 7308
Db 7261 ATGATCCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACTTGAAGAAGC 7320
QY 7309 CTAGATGTCAGATCTACGGGGCTGTACTTCACTTGAAGCACTTGAACCTACCTCAGATC 7368
Db 7321 CTAGATGTCAGATCTACGGGGCTGTACTTCACTTGAAGCACTTGAACCTACCTCAGATC 7380
QY 7369 ATTCAACGACTCCATGSCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7428
Db 7381 ATTCAACGACTCCATGSCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7440
QY 7429 AATAGGTGGCTTCATGCTCAGGAATCTTGGGTACCGCCCTTGGAGTCTGAGACAT 7488
Db 7441 AATAGGTGGCTTCATGCTCAGGAATCTTGGGTACCGCCCTTGGAGTCTGAGACAT 7500
QY 7489 CGGCGCAGAAAGTGTCCGGCTAGGCTACTGTCCAGGGGGGGGCGTCCGCACTTGTGGC 7548
```

```
Db 7501 CGGCGCAGAAAGTGTCCGGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC 7560
QY 7549 AAGTACCTCTTTCAACTGGCGAGTAAGGACCAAGCTCAAACTCAGTCCAAATCCCGGCTGG 7608
Db 7561 AAGTACCTCTTTCAACTGGCGAGTAAGGACCAAGCTCAAACTCAGTCCAAATCCCGGCTGG 7620
QY 7609 TCCAGTTGGATTTATCCAGCTGGTTCGTTGCTGCTTACAGGGGGGAGACATATATCAC 7668
Db 7621 TCCAGTTGGATTTATCCAGCTGGTTCGTTGCTGCTTACAGGGGGGAGACATATATCAC 7680
QY 7669 AGCTGTCTCGTCCCGGACCCCGCTGGTTCATGTGGTCTACTCTACTTCTGTAGGG 7728
Db 7681 AGCTGTCTCGTCCCGGACCCCGCTGGTTCATGTGGTCTACTCTACTTCTGTAGGG 7740
QY 7729 GTAGCATCTATCTACTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7788
Db 7741 GTAGCATCTATCTACTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7800
QY 7789 CCATCTCTTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7848
Db 7801 CCATCTCTTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7849 TTTTCTCTTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7908
Db 7861 TTTTCTCTTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7909 CCTAGTCACGGCTAGCTGTGAAGGTCGTGAGCGGCTGTGACTGCAGAGAGTGTGATAC 7968
Db 7921 CCTAGTCACGGCTAGCTGTGAAGGTCGTGAGCGGCTGTGACTGCAGAGAGTGTGATAC 7980
QY 7969 TGGCTCTCTGCAGATCAAGT 7989
Db 7981 TGGCTCTCTGCAGATCAAGT 8001

RESULT 6
AX036254
LOCUS AX036254
DEFINITION Sequence 3 from Patent EP1043399.
ACCESSION AX036254
VERSION AX036254.1 GI:11225870
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
viral; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Bartenschlager, R.D.
TITLES Hepatitis C virus cell culture system
JOURNAL Patent: EP 1043399-A 3 11-OCT-2000;
BARTENSCHLAGER RALF DR (DE)
FEATURES
source
1. .8001
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:111103"

ORIGIN
Query Match 99.7%; Score 7965.4; DB 6; Length 8001;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7989; Conservative 0; Mismatches 1; Indels 12; Gaps 1;
QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60
Db 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60
QY 61 TCTTCAGCAGAAAGGCTTAGCCATGGCTTAGTATGATGTCGTCGACGCTCCAGGAC 120
Db 61 TCTTCAGCAGAAAGGCTTAGCCATGGCTTAGTATGATGTCGTCGACGCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGGTGAGTACACCCGAAATTCGAC 180
```

Db 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTTCGCGAACCGGTGAGTACACCGGAATTCGCCAG 180  
Qy 181 GACGACCGGGTCTCTTTTGGATCAACCGCTCAATGCTCGAGATTTGGCGTGCCCC 240  
Db 181 GACGACCGGGTCTCTTTTGGATCAACCGCTCAATGCTCGAGATTTGGCGTGCCCC 240  
Qy 241 GCGAGATGCTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTGTGGTACTGCTGATAGG 300  
Db 241 GCGAGATGCTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTGTGGTACTGCTGATAGG 300  
Qy 301 GTGCTTGGAGTGGCCCGGAGGTCTGTAGACCGGTGCAACATGAGCAGCAATTCCTAAAC 360  
Db 301 GTGCTTGGAGTGGCCCGGAGGTCTGTAGACCGGTGCAACATGAGCAGCAATTCCTAAAC 360  
Qy 361 CTCGAAGAAACCAAAA-----GGCGCGCCATGATTGAACGAAGTGAATGTC 408  
Db 361 CTCGAAGAAACCAAAAAGTGAACCAACCGCGCGCCATGATTGAACGAAGTGAATGTC 420  
Qy 409 AGCGAGTTCTCCGCGCGTTCGGGTGGAGAGCTATTTCGGCTATGACTGGGCAACAAGA 468  
Db 421 AGCGAGTTCTCCGCGCGTTCGGGTGGAGAGCTATTTCGGCTATGACTGGGCAACAAGA 480  
Qy 469 CAATCGGCTGCTCATGCGCGCTGTTCCGCTGTGACGCGAGGGCGCCCGGTTCTTT 528  
Db 481 CAATCGGCTGCTCATGCGCGCTGTTCCGCTGTGACGCGAGGGCGCCCGGTTCTTT 540  
Qy 529 TTGTCAAGACCGACCTGTCCGCTGCCCTGAATGAATGCAAGTGCAGACGAGCGCGGCTAT 588  
Db 541 TTGTCAAGACCGACCTGTCCGCTGCCCTGAATGAATGCAAGTGCAGACGAGCGCGGCTAT 600  
Qy 589 CGTGGCTGGCAACGAGCGGCTTCCTTGGCAGCTGTGTGCAAGTGTGCTCACTGAAGCGG 648  
Db 601 CGTGGCTGGCAACGAGCGGCTTCCTTGGCAGCTGTGTGCAAGTGTGCTCACTGAAGCGG 660  
Qy 649 GAAGGACTGCTGCTATTGGCGGAGTGCAGGCGGAGATCTCCTGTCACTCACTGAGCGG 708  
Db 661 GAAGGACTGCTGCTATTGGCGGAGTGCAGGCGGAGATCTCCTGTCACTCACTGAGCGG 720  
Qy 709 CTCCTGCCGAGAAAGTATCCATATGGCTGATGCAATGCGCGGCTGCATACGCTTGATC 768  
Db 721 CTCCTGCCGAGAAAGTATCCATATGGCTGATGCAATGCGCGGCTGCATACGCTTGATC 780  
Qy 769 CGGTACTCTGCCATTCGACCAACGAGCAACATCGCATCGAGCGAGCACTACTCGGA 828  
Db 781 CGGTACTCTGCCATTCGACCAACGAGCAACATCGCATCGAGCGAGCACTACTCGGA 840  
Qy 829 TGGAGCGGCTTGTGATCAGATGATCTGGAGAGAGATCAGGGGCTCGGCGAG 888  
Db 841 TGGAGCGGCTTGTGATCAGATGATCTGGAGAGAGATCAGGGGCTCGGCGAG 900  
Qy 889 CCGAACTGTTCCGCGAGGCTCAAGCGGCGCATGCCGAGCGGAGGATCTCGTGTGACCC 948  
Db 901 CCGAACTGTTCCGCGAGGCTCAAGCGGCGCATGCCGAGCGGAGGATCTCGTGTGACCC 960  
Qy 949 ATGCGATGCTGCTTGGCGAATATCATGFGGAAAAATGCGCGCTTTCTGGATTCATCG 1008  
Db 961 ATGCGATGCTGCTTGGCGAATATCATGFGGAAAAATGCGCGCTTTCTGGATTCATCG 1020  
Qy 1009 ACTGTGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCTTGGCTACCGGTGATA 1068  
Db 1021 ACTGTGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCTTGGCTACCGGTGATA 1080  
Qy 1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCCTGCTTTTACGGTATCGCCG 1128  
Db 1081 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCCTGCTTTTACGGTATCGCCG 1140  
Qy 1129 CTCGGAATTCGAGCGCATCGCTTCTATCGCTTCTGACGAGTTCTTCTGAGTTTAAA 1188  
Db 1141 CTCGGAATTCGAGCGCATCGCTTCTATCGCTTCTGACGAGTTCTTCTGAGTTTAAA 1200  
Qy 1189 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248  
Db 1201 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1260

Qy 1249 AACCTTCTGGCGGAGCGCTTGGAAATAGCGCGGTGCGTTGTCTATATGTTATTT 1308  
Db 1261 AACCTTCTGGCGGAGCGCTTGGAAATAGCGCGGTGCGTTGTCTATATGTTATTT 1320  
Qy 1309 TCCACATATTGCGCTTTTGGCAATGTGAGGGCCCGGAAACCTTGGCTCTCTTTG 1368  
Db 1321 TCCACATATTGCGCTTTTGGCAATGTGAGGGCCCGGAAACCTTGGCTCTCTTTG 1380  
Qy 1369 ACGAGCATTTCTAGGGGTCTTTCCCTCTCGCAAGGAATCAAGGTCTGTGATGTC 1428  
Db 1381 ACGAGCATTTCTAGGGGTCTTTCCCTCTCGCAAGGAATCAAGGTCTGTGATGTC 1440  
Qy 1429 GTGAGGAGAGCATTTCTTCTGGAAGCTTCTTGAAGCAAAACATCTGTAGCAACCT 1488  
Db 1441 GTGAGGAGAGCATTTCTTCTGGAAGCTTCTTGAAGCAAAACATCTGTAGCAACCT 1500  
Qy 1489 TGCAGCGAGCGGAACCCCTCGGCGACAGGTCTCGGCGCAAAAGCCAGGTGA 1548  
Db 1501 TGCAGCGAGCGGAACCCCTCGGCGACAGGTCTCGGCGCAAAAGCCAGGTGA 1560  
Qy 1549 TAAGATACCTGCAAGGCGCAACCCCTGAGTCCACGTTGTGAGTTGATGTTG 1608  
Db 1561 TAAGATACCTGCAAGGCGCAACCCCTGAGTCCACGTTGTGAGTTGATGTTG 1620  
Qy 1609 GAAAGAGTCAATGCTCTCTCAAGCGTATTCAACAAAGGCGCTGAAGATGCCAGAAG 1668  
Db 1621 GAAAGAGTCAATGCTCTCTCAAGCGTATTCAACAAAGGCGCTGAAGATGCCAGAAG 1680  
Qy 1669 GTACCCCTTGTATGGGATCTGATCTGGGCGCTCGGTGCACATGTTTACATGCTTTAG 1728  
Db 1681 GTACCCCTTGTATGGGATCTGATCTGGGCGCTCGGTGCACATGTTTACATGCTTTAG 1740  
Qy 1729 TCGAGGTTAAAAACGTTAGGCCCCCGCAACACGAGGAGCGTGTGTTTCTTTGAAAAA 1788  
Db 1741 TCGAGGTTAAAAACGTTAGGCCCCCGCAACACGAGGAGCGTGTGTTTCTTTGAAAAA 1800  
Qy 1789 CACGATATACCATGGGCGCTATTAGCGCTACTTCCCAACAGAGCGAGGCGCTACTGCG 1848  
Db 1801 CACGATATACCATGGGCGCTATTAGCGCTACTTCCCAACAGAGCGAGGCGCTACTGCG 1860  
Qy 1849 TGCATCATCTAGCTCACAAGCGCGGACAGGACAGGTCGAGGGGAGGTCCAAGTG 1908  
Db 1861 TGCATCATCTAGCTCACAAGCGCGGACAGGACAGGTCGAGGGGAGGTCCAAGTG 1920  
Qy 1909 GTCTCCACCGCAACAATCTTCTCGCGACCTTGCCTCAATGGCGTGTGTTGAGCTGTC 1968  
Db 1921 GTCTCCACCGCAACAATCTTCTCGCGACCTTGCCTCAATGGCGTGTGTTGAGCTGTC 1980  
Qy 1969 TATCATGTCGCGCTCAAGAGACCTTTCGCGGCCCAAGGGGCCAATCACCCAAATGTAC 2028  
Db 1981 TATCATGTCGCGCTCAAGAGACCTTTCGCGGCCCAAGGGGCCAATCACCCAAATGTAC 2040  
Qy 2029 ACCAATGTGACACGAGACCTCTGCTGCTGGCAAGCGCCCCCGGGCGCTTCTCTGACA 2088  
Db 2041 ACCAATGTGACACGAGACCTCTGCTGCTGGCAAGCGCCCCCGGGCGCTTCTCTGACA 2100  
Qy 2089 CCAATGCACTTGGCGAGCTCGGACCTTTACTTGTGTACAGGAGCATGCCGATGTCATCCG 2148  
Db 2101 CCAATGCACTTGGCGAGCTCGGACCTTTACTTGTGTACAGGAGCATGCCGATGTCATCCG 2160  
Qy 2149 GTGCGCGCGGGCGGACAGAGGGAGCTACTCTCCCGAGGCCGCTCTCTACTTGT 2208  
Db 2161 GTGCGCGCGGGCGGACAGAGGGAGCTACTCTCTCCCGAGGCCGCTCTCTACTTGT 2220  
Qy 2209 AAGGGCTTTCGCGCGGTCCACTGCTCTGCCCCCTCGGGGACGCTGTGGGATCTTTCCG 2268  
Db 2221 AAGGGCTTTCGCGCGGTCCACTGCTCTGCCCCCTCGGGGACGCTGTGGGATCTTTCCG 2280  
Qy 2269 GCTGCGGTGTGACCGCGGGGTTCGAGGCGGTGACCTTTGTACCCGTCGAGTCTATG 2328  
Db 2281 GCTGCGGTGTGACCGCGGGGTTCGAGGCGGTGACCTTTGTACCCGTCGAGTCTATG 2340

QY 2329 GAAACCACTATGCGGTCCCGGTCTTACGGCAAACTGTCCTCCCTCCGGCGGTACCGCAG 2388  
DB 2341 GAAACCACTATGCGGTCCCGGTCTTACGGCAAACTGTCCTCCCTCCGGCGGTACCGCAG 2400  
QY 2389 ACATTTCCAGTGGCCCATCTACAGCCCTCTCTGTTAGCGGCAAGACATAAGTGGCG 2448  
DB 2401 ACATTTCCAGTGGCCCATCTACAGCCCTCTCTGTTAGCGGCAAGACATAAGTGGCG 2460  
QY 2449 GCTGGTATGCAAGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCCGTGCGCCACCC 2508  
DB 2461 GCTGGTATGCAAGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCCGTGCGCCACCC 2520  
QY 2509 CTAGGTTTCGGGGCGTATATGTTCTAAGGCACATGTTATCGACCTTAACTACAGAACCGGG 2568  
DB 2521 CTAGGTTTCGGGGCGTATATGTTCTAAGGCACATGTTATCGACCTTAACTACAGAACCGGG 2580  
QY 2569 GTAAGGACCATCACAGCGGTGCCCATCACGTACTCCACCTATGGCAAGTTTCTTGCC 2628  
DB 2581 GTAAGGACCATCACAGCGGTGCCCATCACGTACTCCACCTATGGCAAGTTTCTTGCC 2640  
QY 2629 GACGGTGGTGTCTCTGGGGCGCTATGACATCATATAATATGATGAGTGCCACTCAACT 2688  
DB 2641 GACGGTGGTGTCTCTGGGGCGCTATGACATCATATAATATGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGCTGGAGCG 2748  
DB 2701 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGCTGGAGCG 2760  
QY 2749 CCACTCGTCTGCTGCCACCGCTACGCTCCGGGATCGGTCAACCTGCGCACATCCAAAC 2808  
DB 2761 CCACTCGTCTGCTGCCACCGCTACGCTCCGGGATCGGTCAACCTGCGCACATCCAAAC 2820  
QY 2809 ATCGAGAGGTGGTCTCTCGACACTGGAGAAATCCCTTTATGGCAAGCCATCCCC 2868  
DB 2821 ATCGAGAGGTGGTCTCTCGACACTGGAGAAATCCCTTTATGGCAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGGAGGACCTCATTTCTGCGATTCGAAGAAGAAATGTGAT 2928  
DB 2881 ATCGAGACCATCAAGGGGGGAGGACCTCATTTCTGCGATTCGAAGAAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGGAAGCTGTCCGSCCTCGGACTCAATGCTGTAGCATATTACGGGGCCTT 2988  
DB 2941 GAGCTCGCGGAAGCTGTCCGSCCTCGGACTCAATGCTGTAGCATATTACGGGGCCTT 3000  
QY 2989 GATGTATCGGTATACCAACTAGCGGAGACGTCATTTGCTAGCAACGACGCTCTAATG 3048  
DB 3001 GATGTATCGGTATACCAACTAGCGGAGACGTCATTTGCTAGCAACGACGCTCTAATG 3060  
QY 3049 ACGGGCTTTACGGCGAATTCGACTCAGTGATCGACTGCAATACATGTGTACCCAGACA 3108  
DB 3061 ACGGGCTTTACGGCGAATTCGACTCAGTGATCGACTGCAATACATGTGTACCCAGACA 3120  
QY 3109 GTCGACTTCAGCTCAGCCGACCTTACCATTTAGACGACGACCGGTGCCAAGACGCG 3168  
DB 3121 GTCGACTTCAGCTCAGCCGACCTTACCATTTAGACGACGACCGGTGCCAAGACGCG 3180  
QY 3169 GTGTACGCTCGCAGCGGGGAGGACGACTGTGTAGGGGACGAGATGGGCATTTACAGTTT 3228  
DB 3181 GTGTACGCTCGCAGCGGGGAGGACGACTGTGTAGGGGACGAGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTCCAGGAAAGCGCCCTCGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTAT 3288  
DB 3241 GTGACTCCAGGAAAGCGCCCTCGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTAT 3300  
QY 3289 GACCGGGCTGTGCTGTTGTAAGAGTCAAGCTCAGTCCGCGAGACCTCAGTTAGTTTGGGGCT 3348  
DB 3301 GACCGGGCTGTGCTGTTGTAAGAGTCAAGCTCAGTCCGCGAGACCTCAGTTAGTTTGGGGCT 3360  
QY 3349 TACCTAAACACACAGGGTTGCCGTCTGCCAGACCATCTGGAGTTCTGGGAGAGCGTC 3408  
DB 3361 TACCTAAACACACAGGGTTGCCGTCTGCCAGACCATCTGGAGTTCTGGGAGAGCGTC 3420  
QY 3409 TTACAGGCGCTACCCACATAGAGCGCCCATTTCTTGTCCAGACTAAGACGGCAGGAGAC 3468

DB 3421 TTTCAGGCGCTACCCACATAGACGCCCATTTCTTGTCCGACTAAGCAGCGCAGGAC 3480  
QY 3469 AACTTTCCCTTACCTGGTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACT 3528  
DB 3481 AACTTTCCCTTACCTGGTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACT 3540  
QY 3529 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGACGGGCCA 3588  
DB 3541 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGACGGGCCA 3600  
QY 3589 ACGCCCTTGTCTATAGGTGGAGCCGTTTCAAAACGAGGTTACTACACACACCCCAT 3648  
DB 3601 ACGCCCTTGTCTATAGGTGGAGCCGTTTCAAAACGAGGTTACTACACACACCCCAT 3660  
QY 3649 ACCAAATACATCATGGCATGCTGCGGTGACCTGGAGGTGCTCACGAGACCTGGGTG 3708  
DB 3661 ACCAAATACATCATGGCATGCTGCGGTGACCTGGAGGTGCTCACGAGACCTGGGTG 3720  
QY 3709 CTGTTAGCGGAGTCTTAGCAGCTCTGGCCGGTATTGCTGACACAGGACGCTGTC 3768  
DB 3721 CTGTTAGCGGAGTCTTAGCAGCTCTGGCCGGTATTGCTGACACAGGACGCTGTC 3780  
QY 3769 ATTGTGGGAGCATCATCTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3828  
DB 3781 ATTGTGGGAGCATCATCTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACTCCCTTACATCGAACAGGA 3888  
DB 3841 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACTCCCTTACATCGAACAGGA 3900  
QY 3889 ATGAGCTCGCGCAACATTTCAAACAGAGGAATCGGTTGCTGCAACAGCCACCAAG 3948  
DB 3901 ATGAGCTCGCGCAACATTTCAAACAGAGGAATCGGTTGCTGCAACAGCCACCAAG 3960  
QY 3949 CAAGCGAGGCTGCTGCTCCCGTGGTGAATCAAAGTGGCGGACCTCGAAGCCTTCTGG 4008  
DB 3961 CAAGCGAGGCTGCTGCTCCCGTGGTGAATCAAAGTGGCGGACCTCGAAGCCTTCTGG 4020  
QY 4009 GCGAAGCATATGGAATTTCAATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTG 4068  
DB 4021 GCGAAGCATATGGAATTTCAATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTGCGAACCCCGGATAGCATCATGATGATTCACAGCCTCTATCAGCCGCTC 4128  
DB 4081 CTGCGAACCCCGGATAGCATCATGATGATTCACAGCCTCTATCAGCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTAAATCCTGGGGGATGGGTGGCCGCCCAACTTGTCT 4188  
DB 4141 ACCACCAACATACCTCTCTTTAAATCCTGGGGGATGGGTGGCCGCCCAACTTGTCT 4200  
QY 4189 CTCTCCAGCGCTGCTTCTCTTTGTTAGCGCGCGCATCGCTGGAGCGGCTGTGGAGC 4248  
DB 4201 CTCTCCAGCGCTGCTTCTCTTTGTTAGCGCGCGCATCGCTGGAGCGGCTGTGGAGC 4260  
QY 4249 ATAGGCTTTGGAGGTGCTTGTGATATTTTGGCAGGTTATGGAGAGGGTGGCAGGC 4308  
DB 4261 ATAGGCTTTGGAGGTGCTTGTGATATTTTGGCAGGTTATGGAGAGGGTGGCAGGC 4320  
QY 4309 GCGCTCGTGGCTTTAAGGTTCATGAGCGCGAGATGCCCTCCACCGAGACCTGGTTAAC 4368  
DB 4321 GCGCTCGTGGCTTTAAGGTTCATGAGCGCGAGATGCCCTCCACCGAGACCTGGTTAAC 4380  
QY 4369 CTATCCCTCTATCTCTCCCTGCGCCCTAGTCTCGGGGTCTGTGCGCAGCGATA 4428  
DB 4381 CTATCCCTCTATCTCTCTCCCTGCGCCCTAGTCTCGGGGTCTGTGCGCAGCGATA 4440  
QY 4429 CTGCGTGGGACGTTGGGCGGAGGGGCTGTGCGAGTGAACCGGCTGTAGCG 4488  
DB 4441 CTGCGTGGGACGTTGGGCGGAGGGGCTGTGCGAGTGAACCGGCTGTAGCG 4500  
QY 4489 TTGCTTCCGGGGTAAACACGCTCTCCCGCACCTATGTGCTCAGAGCGGACGCTGCA 4548

Db 4501 TTCCGTCGCGGGTAACCAAGCTCTCCGCCACACTATGTGCCTGAGAGCAAGCTGCA 4560  
QY 4549 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCA 4608  
Db 4561 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCA 4620  
QY 4609 CAGTGGATCAACAGAGGAGTGTCTCCACGCGCATCTCCGGCTCGTGCTTAAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACAGAGGAGTGTCTCCACGCGCATCTCCGGCTCGTGCTTAAGAGATGTTGG 4680  
QY 4669 GATTGGATATGACCGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAAAGTCTCTGCG 4728  
Db 4681 GATTGGATATGACCGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAAAGTCTCTGCG 4740  
QY 4729 CGATTGCGGGAGTCCCTCTTCTCATGTCTCAAGTGGGTACAAAGGAGTCTGGCGGGC 4788  
Db 4741 CGATTGCGGGAGTCCCTCTTCTCATGTCTCAAGTGGGTACAAAGGAGTCTGGCGGGC 4800  
QY 4789 GACGGCATCATGCAAAACCACTGCCCCATGTGGAGCACAGATCACCGGACATGTGAAAAA 4848  
Db 4801 GACGGCATCATGCAAAACCACTGCCCCATGTGGAGCACAGATCACCGGACATGTGAAAAA 4860  
QY 4849 GGTTCATGAGATCTGCGGGCTTAGGACCTGTAGTATACAGTGTGCATGGAATTCGCC 4908  
Db 4861 GGTTCATGAGGATCTGCGGGCTTAGGACCTGTAGTATACAGTGTGCATGGAATTCGCC 4920  
QY 4909 ATTAAACGGGTATACCAACGCGGCCCTGACGCGCTCCCGGCGCCAAATTTCTAGGGG 4968  
Db 4921 ATTAAACGGGTATACCAACGCGGCCCTGACGCGCTCCCGGCGCCAAATTTCTAGGGG 4980  
QY 4969 CTGTGGCGGGTGGCTGTGAGAGTACGTGGAGGTTACGCGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGGTGGCTGTGAGAGTACGTGGAGGTTACGCGGGTGGGGATTTCCACTAC 5040  
QY 5029 GTGACGGGCATGACCACTGACAAAGTAAAGTGCCTGTCTAGGTTCCCGCCCCGAAATTC 5088  
Db 5041 GTGACGGGCATGACCACTGACAAAGTAAAGTGCCTGTCTAGGTTCCCGCCCCGAAATTC 5100  
QY 5089 TTCAAGAGTGGATGGGGTGGTGGTGCACAGGTACGCTCCAGCGTGCAAAACCCCTCTTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGTGGTGCACAGGTACGCTCCAGCGTGCAAAACCCCTCTTA 5160  
QY 5149 CGGAGGAGGTCAATTCCTGGTGGGCTCAATCAATACTGGTTGGGTACAGCTCCA 5208  
Db 5161 CGGAGGAGGTCAATTCCTGGTGGGCTCAATCAATACTGGTTGGGTACAGCTCCA 5220  
QY 5209 TGCAGCCGCAACCGGACGTAGAGTGTCTCTTCATGTCTACGACCCCTCCCAAT 5268  
Db 5221 TGCAGCCGCAACCGGACGTAGAGTGTCTCTTCATGTCTACGACCCCTCCCAAT 5280  
QY 5269 ACGCGGAGACGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCA 5328  
Db 5281 ACGCGGAGACGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGTACGACGTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGTCACTGCTC 5388  
Db 5341 TCAGTACGACGTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGTCACTGCTC 5400  
QY 5389 CCGGACGCTGACCTATCGAGGCCAACCTCTGTGGCGGACGAGATGGCGGGAAATTC 5448  
Db 5401 CCGGACGCTGACCTATCGAGGCCAACCTCTGTGGCGGACGAGATGGCGGGAAATTC 5460  
QY 5449 ACCCGCTGGAGTCAGAAAAATAGGTAGTAAATTTGGACTCTTTGAGCGGCTCCAGCG 5508  
Db 5461 ACCCGCTGGAGTCAGAAAAATAGGTAGTAAATTTGGACTCTTTGAGCGGCTCCAGCG 5520  
QY 5509 GAGGAGATGAGAGGAGTATCCGTTCCGCGGAGATCTCTGGGAGTCCAGGAAATTC 5568  
Db 5521 GAGGAGATGAGAGGAGTATCCGTTCCGCGGAGATCTCTGGGAGTCCAGGAAATTC 5580  
QY 5569 CCTGAGCGATGCCATATGGGACCGCCGANTTAAACCTTCACTGTTAGATCTCTGG 5628  
Db 5581 CCTGAGCGATGCCATATGGGACCGCCGANTTAAACCTTCACTGTTAGATCTCTGG 5640

QY 5629 AAGGACCCG3ACTACGTCCCTCCAGTGTGTACACGGGTGTCTCAATGCGCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCCG3ACTACGTCCCTCCAGTGTGTACACGGGTGTCTCAATGCGCGCTGCCAAGGCC 5700  
QY 5689 CCTCGATACACACTCCACGAGGAGAGGAGGTTGCTCTGTAGATCTTACCGTGTCT 5748  
Db 5701 CCTCGATACACACTCCACGAGGAGAGGAGGTTGCTCTGTAGATCTTACCGTGTCT 5760  
QY 5749 TCTGCTTTGGG3AGTCTCGCCACAAAGACCTTTCCGAGCTCCGAATCTCGCGCCCTCGAC 5808  
Db 5761 TCTGCTTTGGG3AGTCTCGCCACAAAGACCTTTCCGAGCTCCGAATCTCGCGCCCTCGAC 5820  
QY 5809 AGCGGACCGGCAACGGGCTCTCTGACAGGCCCTCCGACGCGGAGCGCGGATCCGAC 5868  
Db 5821 AGCGGACCGGCAACGGGCTCTCTGACAGGCCCTCCGACGCGGAGCGCGGATCCGAC 5880  
QY 5869 GTTGAGTCTACTCTCTCATGCCCTTCCAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTACTCTCTCATGCCCTTCCAGGGGAGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGTCTTGGTCTACCGTAAAGGAGGAGTGTAGAGGAGTGTCTGTCTGTCTCGATG 5988  
Db 5941 GACGGTCTTGGTCTACCGTAAAGGAGGAGTGTAGAGGAGTGTCTGTCTGTCTCGATG 6000  
QY 5989 TCCTACATGACGAGCGGCCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGCTGCC 6048  
Db 6001 TCCTACATGACGAGCGGCCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGCTGCC 6060  
QY 6049 ATCAATGACATGAGCACTCTTTGCTCCGTCCACCAAACTTTGGTCTATGTCTACAACT 6108  
Db 6061 ATCAATGACATGAGCACTCTTTGCTCCGTCCACCAAACTTTGGTCTATGTCTACAACT 6120  
QY 6109 CGCAGCGAAGCTTGGCGGACAGAAAGTCACTTTGACAGACTGCGAGTCTCGAGCAG 6168  
Db 6121 CGCAGCGAAGCTTGGCGGACAGAAAGTCACTTTGACAGACTGCGAGTCTCGAGCAG 6180  
QY 6169 CACTACCGGAGTGTCTCAGGAGATGAGCGGAGGCTCCAGTTAAGGCTAAATTT 6228  
Db 6181 CACTACCGGAGTGTCTCAGGAGATGAGCGGAGGCTCCAGTTAAGGCTAAATTT 6240  
QY 6229 CTATCCGTGAGGAGGCTGTAACTGACGCGCCCACTTCGCGCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGAGGAGGCTGTAACTGACGCGCCCACTTCGCGCAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAGGACGTCCGGAACCTATCCAGCAAGGCGGTTAACCACTCCGCTCCGTG 6348  
Db 6301 TATGGGCAAGGACGTCCGGAACCTATCCAGCAAGGCGGTTAACCACTCCGCTCCGTG 6360  
QY 6349 TGGAGGACTGTCTGAGGACACTGAGACACCAATTTGACACCACTCATCGGCAAAAT 6408  
Db 6361 TGGAGGACTGTCTGAGGACACTGAGACACCAATTTGACACCACTCATCGGCAAAAT 6420  
QY 6409 GAGTTTTCTCGCTCCAAACAGAGAGGCGGCGCAAGCGCTGCGCTTATCGTATTC 6468  
Db 6421 GAGTTTTCTCGCTCCAAACAGAGAGGCGGCGCAAGCGCTGCGCTTATCGTATTC 6480  
QY 6469 CCAGATTTGGGGTTCGTGTGCGAGAAATGAGGCTTTTACGATGGGTCTCCACCTC 6528  
Db 6481 CCAGATTTGGGGTTCGTGTGCGAGAAATGAGGCTTTTACGATGGGTCTCCACCTC 6540  
QY 6529 CCTCAGCGCTGATGGGCTCTTCATACGATTCGAATCTCTCTGACAGCGGCTCGAG 6588  
Db 6541 CCTCAGCGCTGATGGGCTCTTCATACGATTCGAATCTCTCTGACAGCGGCTCGAG 6600  
QY 6589 TTCTGTGTAATGCTGGAAAGCGAAGAAATGCGCTTATGGGCTTCGATATGACACCGC 6648  
Db 6601 TTCTGTGTAATGCTGGAAAGCGAAGAAATGCGCTTATGGGCTTCGATATGACACCGC 6660  
QY 6649 TGTGTTGACTCAACGGTCACTGAGATGACATCGGTGTTGAGGAGTCAATCTACCAATGT 6708  
Db 6661 TGTGTTGACTCAACGGTCACTGAGATGACATCGGTGTTGAGGAGTCAATCTACCAATGT 6720



6709	QY	TGTGACTTTGGCCCCGAAAGCAGACAGAGCGGCTTTACATC	6766
6721	Db	TGTGACTTTGGCCCCGAAAGCAGACAGAGCGGCTTTACATC	6780
6769	QY	GGGGCCCCCTGACTAAATTTCTAAAGGGCAGACTCGGGCTATCGCCGGTGGCCGCGGAGC	6828
6781	Db	GGGGCCCCCTGACTAAATTTCTAAAGGGCAGACTCGGGCTATCGCCGGTGGCCGCGGAGC	6840
6829	QY	GGTGTACTGACCAACAGCTCGCGTAATAACCTCACATGTTTACTTTGAAGCGCGCTCGGGCC	6888
6841	Db	GGTGTACTGACCAACAGCTCGCGTAATAACCTCACATGTTTACTTTGAAGCGCGCTCGGGCC	6900
6889	QY	TGTGAGCTGCGAAGCTTCCAGACTGCGACGATGCTGTATGCGGAGACGACTTGTGCTT	6948
6901	Db	TGTGAGCTGCGAAGCTTCCAGACTGCGACGATGCTGTATGCGGAGACGACTTGTGCTT	6960
6949	QY	ATCTGTGAAGCGGGGACCCAAAGGACGAGCGGAGCCTACGGGCTTTCACGAGGCT	7008
6961	Db	ATCTGTGAAGCGGGGACCCAAAGGACGAGCGGAGCCTACGGGCTTTCACGAGGCT	7020
7009	QY	ATGACTAGATACTCTGCCCCCTTGGGGACCCGCCCAAACAGAAATACGACTTGGAGTTG	7068
7021	Db	ATGACTAGATACTCTGCCCCCTTGGGGACCCGCCCAAACAGAAATACGACTTGGAGTTG	7080
7069	QY	ATAACTATGCTCTCCAAATGTGTCAGTCGCGGACGATGCATCTGGCAAAAGGTGTATC	7128
7081	Db	ATAACTATGCTCTCCAAATGTGTCAGTCGCGGACGATGCATCTGGCAAAAGGTGTATC	7140
7129	QY	TATCTACCCGTCAGCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC	7188
7141	Db	TATCTACCCGTCAGCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC	7200
7189	QY	ACTCCAGTCAATTCCTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGG	7248
7201	Db	ACTCCAGTCAATTCCTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGG	7260
7249	QY	ATGATCCTGATGACTCATTTCTTCTCCACTTTCTAGCTCAGGAACAATTTGAAAAGCC	7308
7261	Db	ATGATCCTGATGACTCATTTCTTCTCCACTTTCTAGCTCAGGAACAATTTGAAAAGCC	7320
7309	QY	CTAGATTGTCAGATCTACGGGCTGTATCTCCATTCAGCCACTTGGACCTACCTCAGATC	7368
7321	Db	CTAGATTGTCAGATCTACGGGCTGTATCTCCATTCAGCCACTTGGACCTACCTCAGATC	7380
7369	QY	ATTCAACGACTCCATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC	7428
7381	Db	ATTCAACGACTCCATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC	7440
7429	QY	AATAGGTGGCTTTCATGCCCTCAGAAACTTTGGGTACCGCCCTTGCAGTCTGGAGACAT	7488
7441	Db	AATAGGTGGCTTTCATGCCCTCAGAAACTTTGGGTACCGCCCTTGCAGTCTGGAGACAT	7500
7489	QY	CGGGCCAGAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC	7548
7501	Db	CGGGCCAGAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC	7560
7549	QY	AAGTACCTCTTCAACTGGCAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGC	7608
7561	Db	AAGTACCTCTTCAACTGGCAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGC	7620
7609	QY	TCCCAGTTGGATTTATCAGCTGGTTCGTTGCTGGTTACAGCGGGGGAGACATATATCAC	7668
7621	Db	TCCCAGTTGGATTTATCAGCTGGTTCGTTGCTGGTTACAGCGGGGGAGACATATATCAC	7680
7669	QY	AGCCTGTCTGTGCGGACCCCGCTGGTTCATGTGGTGGCTACTCTTCTGTGTAGGG	7728
7681	Db	AGCCTGTCTGTGCGGACCCCGCTGGTTCATGTGGTGGCTACTCTTCTGTGTAGGG	7740
7729	QY	GTAGGCATCTATCTACTCCCCAAACCGATGAACGGGACCTTAAACACTCCAGGCCAATAGG	7788
7741	Db	GTAGGCATCTATCTACTCCCCAAACCGATGAACGGGAGCTTAAACACTCCAGGCCAATAGG	7800
7789	QY	CCATCCCTGTTTTTTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	7848

[illegible]



```
/db_xref="GI:5441841"
/translation="MSTNPKPQKTKRNTNGRAMTEODQLHAGSPAAWVERLFGVDWA
QOTIGCSAAAPRLSAOGRPVLFVKTDLSGALNELQDEARLSWLTATGVPAAVLVD
VTEARGLLGEVPGQDLISHLAPAEKYSIMADAMRLHLTDPATCFPHQAKHRI
EAPRTMAGLVQDDLBEBHQGLPAELFARKKAMPDGEDLVVYTHGDACLPLNMVE
NRFSGFIDCGRLGVADRYQDIALATRDIAELGGEWADRFVLVYGLIAAPDSQRIAFY
RLDLDF"
1202..1812
/misc_feature
/feature="internal ribosome entry site (IRES)"
1813..7770
/codon_start=1
/product="non-structural polyprotein"
/protein_id="CAB46917.1"
/db_xref="GI:5441842"
/translation="MAPITAYSQOTRGLLGCITSLTGRDRNOVEGEVQVSTAQSP
LATVNGVCWTVYHAGSKTLGAPKGPITQMYTNVDQLVQWAPPGARSLTPCTCGS
SDLYLVTHADVIVRRGRDGRSLSPRPVSLKSGSGGLLCPSHGAVGIFRAAVC
TRGAKAVLPVFMESMETMRSPVETDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAA
YAAQYKVLVLPNSVATLGFAYMSKAHGDPINRITGVRTITGAPITYSTYKFLA
DGGSGGAYDIIICDECHSTDSTILGITGVLDQAEATAGARLVLTATPPGSVTPH
PNIEVALSSTGEIPFYCKAIPETIKGRHLIFCHSKKCDLAAKLSGLGLNAVY
YRGLDVSIPSGDIVVATDALMTGFTGDPDSVDNCTQVQTFWFSLDPTFIETI
TYPQDASRSRGRTRGRNGIYRVTTPGERPSPFSSVLCBECYDAGCAWELTPA
ETSVRLAYLTNPLPVCODHLEFWESVFTGLTHIDAHFSLQTKQAGDNFPYLVAQA
TVCARAAPPSPSQMCKLLRLKPTLHGPTPLLYRLGAVQNEVTTTHPIIKYIMACM
SADLEVSTVTLVGVGLAAALAAVCLATGTSVIVGRIILSGKPAIIPREVLREFDE
MEECASHLPYTEQGMOLAEOKAKAIGLLOATKQABAAAPVVEKWTLEAFWAKM
WNFTSIGIYLAGISLTPGNPAIASIEMFTASITSLPTTQHTLLFNILGQWAAQLAP
SAAAFVAGIAGTAGVSGIGLKVLVDILAGYAGVAGALVAFKVMGEMSTEDLVN
LPLALLSGAVGVGCAAILRRHVPGEGAVQWNNRLIAFASRGNHVSPTHYPESD
AAAVQTLLSLATQLKRLHINEDCSTPCSGWLRDVMDWICTVLTDFKWLQS
KLLPLPGVPFPGCGYKGVMGDGMQITPCGCAQITGHVKNGSMRI VGPRTCSNT
WHGTFPIINAYTTPCTGTPAPNSRALWRVAEEVETRVYGVHYVTGMTDVNKP
COVPAFEFTVDGVRHLRYAPACKPILLREVTFLVGLNQVLVGSQLPCEBEPDVL
TSMLTDSHITAEAKRLARGSPSLASSASQSLAPSLKATCTTRHSDPADLIEA
NLMLQKMGNTIVSEKNVVIILDSFEPLQAEDEREVSPAILRRSRKFPFRAMPI
WARDPNPLLESKVDYPPVHVHGFPLPAKAPPIPPRRKRTVLVSESTVSALA
ELATKTSSESSAVDSGTASDPQSDSDAGSDVBSYSPMPLEGEQDPLSDG
SWSTVSEASBDVCCMSYTWGALITPCAABETKLPIINALSNLLRHLNVLVATTS
RSASLRQKVTFDRLQVLDHVRDLVKEMKAKASTVKAKLLSVBEACKLTPPHSARSK
FGYKADVRNLSKAVNHIRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGRKPKR
LIVPDLGVRCERKALYDVVSTLPQAVMGSSYGFQYSPGQVBFVLNMAKKCPMG
FAYTRCFDSTVENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKQNC
GYRCSAGVLVITSGNLTCLYKAAACRAAKLQDCQTLVCGDLDLVVICBSAGTQED
EASLARATEAMTRYSAAPGDPKPEYDLELITSCSNVSVAHDSGRKRVYLTDPPT
PLARAWETARWTPVNSLWNIIMVAPTLWARMILMTHFFSILLAQOLEKALDQIY
GACYSIRPLDLPOILQRLHGLSAPLSHSYSGEINRVASCLRLKGLVPLRVRHRAS
VRALLSQGEARATCGKYLNNWAVRTKLKLTPIPAASQLDLSWFWAGYSYGGDIYHSL
SRAPRPFMMWCLLLLSVGVGIYLLPNR"
1813..3708
/gene="NS3"
/product="NS3 proteinase/helicase"
3709..3870
/gene="NS4A"
/product="NS3/4A proteinase cofactor"
3871..4653
/gene="NS4B"
/product="NS4B protein"
4654..5994
/gene="NS5A"
/product="NS5A phosphoprotein"
5995..7767
/gene="NS5B"
/product="NS5B RNA dependant RNA polymerase"
1813..3708
/gene="NS3"
3709..3870
/gene="NS4A"
3871..4653
/gene="NS4B"
4654..5994
/gene="NS5A"
5995..7767
/gene="NS5B"

mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide
gene
gene
gene
gene
gene
gene
gene
ATGCGCATGCTGCTTGTCCCGAATATCATGTGGGAAATGGCCGCTTTTCTTGATTCATCG 1008
```

```
/gene="NS5B"
7771..8001
3'UTR
ORIGIN
Query Match 99.7%; Score 7965.4; DB 12; Length 8001;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 12; Gaps 1;
QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCAATAGATCACTCCCTGTGAGAACTACTG 60
DB 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCAATAGATCACTCCCTGTGAGAACTACTG 60
QY 61 TCTTCACGCAAGAAACGCTCTAGCCATGCGCTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120
DB 61 TCTTCACGCAAGAAACGCTCTAGCCATGCGCTTAGTATAGTGTCTGTCGAGCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCAG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCAG 180
QY 181 GACGACCGGGTCTCTTTTGGATCAACCGGCTCAATGCTGAGATTTGGGGCGTCCGCC 240
DB 181 GACGACCGGGTCTCTTTTGGATCAACCGGCTCAATGCTGAGATTTGGGGCGTCCGCC 240
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTTGTGGTACTGCTGATAGG 300
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTTGTGGTACTGCTGATAGG 300
QY 301 GTGCTTGGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGACGACGATCTCTAAAC 360
DB 301 GTGCTTGGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGACGACGATCTCTAAAC 360
QY 361 CTCAGAGAAACCAAA-----GGGCGGCCATGATTGAAACAAGATGATTCG 408
DB 361 CTCAGAGAAACCAAAACCAACCAACGCGCGCGCCATGATTGAAACAAGATGATTCG 420
QY 409 ACGCAGGTTCTCCGCGCGCTTGGGTGAGAGGCTATTTCGGCTATGATCTGCGCAACAGA 468
DB 421 ACGCAGGTTCTCCGCGCGCTTGGGTGAGAGGCTATTTCGGCTATGATCTGCGCAACAGA 480
QY 469 CAATCGGCTGTCTGATGCGCCGCTGTTCGGCTGTGAGCGCAGGGCGCGCGTCTTT 528
DB 481 CAATCGGCTGTCTGATGCGCCGCTGTTCGGCTGTGAGCGCAGGGCGCGCGTCTTT 540
QY 529 TTGTCAGACCGCCTGTCCGCTGCGCTGATGATGATGATGATGATGATGATGATGATGAT 588
DB 541 TTGTCAGACCGCCTGTCCGCTGCGCTGATGATGATGATGATGATGATGATGATGATGAT 600
QY 589 CGTGGCTGGCCACGACGCGCGCTTCTTTCGCGAGCTGTGTCGACGTTCTCACTGAAGCGG 648
DB 601 CGTGGCTGGCCACGACGCGCGCTTCTTTCGCGAGCTGTGTCGACGTTCTCACTGAAGCGG 660
QY 649 GAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCATCTCACCTTG 708
DB 661 GAAGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCATCTCACCTTG 720
QY 709 CTCCTCGCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGATGATGATGATGATGAT 768
DB 721 CTCCTCGCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGATGATGATGATGATGAT 780
QY 769 CGGTACTCTGCCATTTCGACCAACAGCGAAACATCCGCTGAGCGAGCAGTACTCGGA 828
DB 781 CGGTACTCTGCCATTTCGACCAACAGCGAAACATCCGCTGAGCGAGCAGTACTCGGA 840
QY 829 TGGAGCGGCTTGTGTCGATCAGGATGATCTGACGAGAGATCAGGGGCTCCGCGCAG 888
DB 841 TGGAGCGGCTTGTGTCGATCAGGATGATCTGACGAGAGATCAGGGGCTCCGCGCAG 900
QY 889 CCGAATCTTTCGCGAGCTCAAGCGCGCATGCCCGGCGGAGGATCTCGTCTGATGCC 948
DB 901 CCGAATCTTTCGCGAGCTCAAGCGCGCATGCCCGGCGGAGGATCTCGTCTGATGCC 960
QY 949 ATGCGCATGCTGCTTGTCCCGAATATCATGTGGGAAATGGCCGCTTTTCTTGATTCATCG 1008
```

Db 961 ATGGCGATGCCGCTTGGCGAATATATGTTGGAAATATGGCCGCTTTCTGTGATTCATCG 1020  
QY 1009 ACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGTGATA 1068  
Db 1021 ACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCCGTGATA 1080  
QY 1069 TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCCTGTGCTGTATCAGGTATCGCGG 1128  
Db 1081 TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCCTGTGCTGTATCAGGTATCGCGG 1140  
QY 1129 CTCGCCATTCGACGCGATCGGCTTCTATCGGCTTCTTACGAGTTCCTTCTGAGTTTAAA 1188  
Db 1141 CTCGCCATTCGACGCGATCGGCTTCTATCGGCTTCTTACGAGTTCCTTCTGAGTTTAAA 1200  
QY 1189 CAGACCAACAACGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248  
Db 1201 CAGACCAACAACGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1260  
QY 1249 AAGCTTACTGGCGGAAGCCGCTTGGAAATAGGCGCGGTGTGGTGTCTATATGTTATTT 1308  
Db 1261 AAGCTTACTGGCGGAAGCCGCTTGGAAATAGGCGCGGTGTGGTGTCTATATGTTATTT 1320  
QY 1309 TCCACCATATGCCGCTTTTGGCAATGAGGCGCCGAAACCTGGCCCTGTCTCTTTG 1368  
Db 1321 TCCACCATATGCCGCTTTTGGCAATGAGGCGCCGAAACCTGGCCCTGTCTCTTTG 1380  
QY 1369 AGGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGTGAATGTC 1428  
Db 1381 ACGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGTGAATGTC 1440  
QY 1429 GTGAAGGAAGCAGTCTCTGTGAAGTCTTGAAGAACAAACAAACGCTGTAGCGACCCCTT 1488  
Db 1441 GTGAAGGAAGCAGTCTCTGTGAAGTCTTGAAGAACAAACAAACGCTGTAGCGACCCCTT 1500  
QY 1489 TCGAGCAGCGGAACCCCACTGTGCGACAGTGCTCTGGGCCAAAGCCACGCTGA 1548  
Db 1501 TCGAGCAGCGGAACCCCACTGTGCGACAGTGCTCTGGGCCAAAGCCACGCTGA 1560  
QY 1549 TAAGATACACCTGCAAGGCGGCACAAACCCAGTGGCCAGTGTGTGAGTTGGATGTTG 1608  
Db 1561 TAAGATACACCTGCAAGGCGGCACAAACCCAGTGGCCAGTGTGTGAGTTGGATGTTG 1620  
QY 1609 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1668  
Db 1621 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1680  
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCTCTGGTGACATGCTTTACATGTGTTAG 1728  
Db 1681 GTACCCCATTTGATGGATCTGATCTGGGCTCTGGTGACATGCTTTACATGTGTTAG 1740  
QY 1729 TCGAGTTAAAAACGCTTAGGCCCCCGAACACAGCGGACGCTGTTTCTTTGAAAAA 1788  
Db 1741 TCGAGTTAAAAACGCTTAGGCCCCCGAACACAGCGGACGCTGTTTCTTTGAAAAA 1800  
QY 1789 CAGGATAATACCATGGCGCTATTAGCGCTACTCCACAGACGCGAGCCCTACTTGGC 1848  
Db 1801 CAGGATAATACCATGGCGCTATTAGCGCTACTCCACAGACGCGAGCCCTACTTGGC 1860  
QY 1849 TGCATCATCACTAGCTCTCAGCGCGGACAGGAACAGTCCAGCGGGAGGTCCCAAGTG 1908  
Db 1861 TGCATCATCACTAGCTCTCAGCGCGGACAGGAACAGTCCAGCGGGAGGTCCCAAGTG 1920  
QY 1909 GTCTCCACCGCAACAATCTTTCTCTGGGACCTGGCTCAATGGCGTGTGTTGGACTGTC 1968  
Db 1921 GTCTCCACCGCAACAATCTTTCTCTGGGACCTGGCTCAATGGCGTGTGTTGGACTGTC 1980  
QY 1969 TATCATGTGCGCGCTCAAGAACCTTTGCGGCCCAAGGGCCCAATCACCCAAATGTAC 2028  
Db 1981 TATCATGTGCGCGCTCAAGAACCTTTGCGGCCCAAGGGCCCAATCACCCAAATGTAC 2040  
QY 2029 ACCAATGTGACAGGACCTCGTGGCTGGCAGCGCCCGCGGCGGCTTCTCTTGACA 2088

Db 2041 ACCAATGTGACAGGACCTCGTGGCTGGCAAGCGCCCGGCGGCTTCTCTTGACA 2100  
QY 2089 COATGCACCTGCGGAGCTCGGACCTTTACTTTGTGTCAAGAGCATGCGGATGTCATTCG 2148  
Db 2101 COATGCACCTGCGGAGCTCGGACCTTTACTTTGTGTCAAGAGCATGCGGATGTCATTCG 2160  
QY 2149 GTGCGCGCGGCGGCGACAGCAGGCGGAGCTTACTTCTCCCGAGCGCGCTCTCTACTTTG 2208  
Db 2161 GTGCGCGCGGCGGCGACAGCAGGCGGAGCTTACTTCTCCCGAGCGCGCTCTCTACTTTG 2220  
QY 2209 AAGGGCTCTTCTGCGGCGGTCCTACTGTCTGCGCCCTCGGGGACGCTGTGGGCACTTTTCG 2268  
Db 2221 AAGGGCTCTTCTGCGGCGGTCCTACTGTCTGCGCCCTCGGGGACGCTGTGGGCACTTTTCG 2280  
QY 2269 GTGCGCGGTGTGACCCGAGGCGGTTCGGAAGGCGGTGGACTTTGTATCCCGTCGAGTCTATG 2328  
Db 2281 GCTGCGGTGTGACCCGAGGCGGTTCGGAAGGCGGTGGACTTTGTATCCCGTCGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCCGCGCTACCGCAG 2388  
Db 2341 GAAACCACTATGCGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCCGCGCTACCGCAG 2400  
QY 2389 ACATTTCCAGGTGCGCCATCTACACGCCCCCTACTGTGTAGCGGCAAGACATAAGGTGCG 2448  
Db 2401 ACATTTCCAGGTGCGCCATCTACACGCCCCCTACTGTGTAGCGGCAAGACATAAGGTGCG 2460  
QY 2449 GTCGCTATGCAAGCCCAAGGTTAAGGTGCTTGTCTGTAAACCCGTCGTCGCGCCAC 2508  
Db 2461 GCTGCGTATGCAAGCCCAAGGTTAAGGTGCTTGTCTGTAAACCCGTCGTCGCGCCAC 2520  
QY 2509 CTAGTTTTCGCGGCTTATATGCTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2568  
Db 2521 CTAGTTTTCGCGGCTTATATGCTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2580  
QY 2569 GTAAGGACCATCACACGCGTCCCGCTACAGCTACTCCACCTATGGCAAGTTCTTTGCC 2628  
Db 2581 GTAAGGACCATCACACGCGTCCCGCTACAGCTACTCCACCTATGGCAAGTTCTTTGCC 2640  
QY 2629 GACGTTGTTGCTCTGGGCGGCTTATGACATCATTAATGTGATGAGTGCACTCAACT 2688  
Db 2641 GACGTTGTTGCTCTGGGCGGCTTATGACATCATTAATGTGATGAGTGCACTCAACT 2700  
QY 2689 GACTCGACACTTCTGGGATCGGACAGTCTGTGACCAAGCGAGACGCTGGAGG 2748  
Db 2701 GACTCGACACTTCTGGGATCGGACAGTCTGTGACCAAGCGAGACGCTGGAGG 2760  
QY 2749 CGACTCGTGTCTGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCACATCCAAAC 2808  
Db 2761 CGACTCGTGTCTGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCACATCCAAAC 2820  
QY 2809 ATCGAGGAGTGTCTGTCTGAGCACTGGAGAAAATCCCTTTTATGGGAAAGCATCCCC 2868  
Db 2821 ATCGAGGAGTGTCTGTCTGAGCACTGGAGAAAATCCCTTTTATGGGAAAGCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGCGGAGGACCTTCTGCAATTTCTGCAATTTCCAAAGAAATGTGAT 2928  
Db 2881 ATCGAGACCATCAAGGGGCGGAGGACCTTCTGCAATTTCTGCAATTTCCAAAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGCGAAGCTGTCCGCGCTCGGACTCAATGTCTGTAGCATATTACCGGGGCTT 2988  
Db 2941 GAGCTCGCGCGAAGCTGTCCGCGCTCGGACTCAATGTCTGTAGCATATTACCGGGGCTT 3000  
QY 2989 GATGTATCGCTCATACCAACTAGCGGAGAGCTTATGTCTGTAGCAACGAGCGCTCTAATG 3048  
Db 3001 GATGTATCGCTCATACCAACTAGCGGAGAGCTTATGTCTGTAGCAACGAGCGCTCTAATG 3060  
QY 3049 ACGGGCTTTACCGGCAATTTGCACTCAGTGATCGCTGCAATACATGTCTCACCAGACA 3108  
Db 3061 ACGGGCTTTACCGGCAATTTGCACTCAGTGATCGCTGCAATACATGTCTCACCAGACA 3120  
QY 3109 GTGACTTTCAGCTGAGCCGACCTTTCACCATTTGAGACGACCGCTGCGCAACAAGCGG 3168  
Db 3121 GTGACTTTCAGCTGAGCCCGACCTTTCACCATTTGAGACGACCGCTGCGCAACAAGCGG 3180

QY	3169	GTGTCAAGCTCGACGGCGAGGACAGTGGTAGGGCAGATGGGCATTTACAGGTTT	3228
Db	3181	GTGTCAAGCTCGACGGCGAGGACAGTGGTAGGGCAGATGGGCATTTACAGGTTT	3240
QY	3229	GTGACTCCAGGAGAACGGCCCTCGGGCATGTCGATTCTCGGTTCTGTGCGAGTGCTAT	3288
Db	3241	GTGACTCCAGGAGAACGGCCCTCGGGCATGTCGATTCTCGGTTCTGTGCGAGTGCTAT	3300
QY	3289	GACGGGCTGTGCTTGTGTACAGACTCAGCCCGCCGAGACCTCATGTTAGGTTGCGGCT	3348
Db	3301	GACGGGCTGTGCTTGTGTGTACAGACTCAGCCCGCCGAGACCTCATGTTAGGTTGCGGCT	3360
QY	3349	TACCTTAAACACACAGAGGTGGCCGTCTGCGAGGACCATCTGGAGTTCTGGAGAGCGTC	3408
Db	3361	TACCTTAAACACACAGAGGTGGCCGTCTGCGAGGACCATCTGGAGTTCTGGAGAGCGTC	3420
QY	3409	TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGCAGAGAC	3468
Db	3421	TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGCAGAGAC	3480
QY	3469	AACCTCCCTACCTGGTAGCATACAGAGTACGGTGTGGCGCAGGGCTCAGGCTCCACT	3528
Db	3481	AACCTCCCTACCTGGTAGCATACAGAGTACGGTGTGGCGCAGGGCTCAGGCTCCACT	3540
QY	3529	CCATCGTGGGACCAAACTGTGAAGTGTCTCATAGGCTAAAGCCTACGCTGCACGGGCA	3588
Db	3541	CCATCGTGGGACCAAACTGTGAAGTGTCTCATAGGCTAAAGCCTACGCTGCACGGGCA	3600
QY	3589	ACGCCCTGTGTATPAGGCTGGAGCGTTTCAAAACGAGGTTTACTACACACACCCCAT	3648
Db	3601	ACGCCCTGTGTATPAGGCTGGAGCGTTTCAAAACGAGGTTTACTACACACACCCCAT	3660
QY	3649	ACCAATACATCATGGCATGTCGCTGACCTGGAGGTGTCACAGACACCTGGGTG	3708
Db	3661	ACCAATACATCATGGCATGTCGCTGACCTGGAGGTGTCACAGACACCTGGGTG	3720
QY	3709	CTGCTAGGCGGAGTCTTAGAGCTCTGGCCGCTATTTCCTGACAAACAGGAGGTTGTC	3768
Db	3721	CTGCTAGGCGGAGTCTTAGAGCTCTGGCCGCTATTTCCTGACAAACAGGAGGTTGTC	3780
QY	3769	ATTGTGGGAGGATCATTTTGTCCGAAAGCCGGCCATCATTTCCCGACAGGAAAGTCCTT	3828
Db	3781	ATTGTGGGAGGATCATTTTGTCCGAAAGCCGGCCATCATTTCCCGACAGGAAAGTCCTT	3840
QY	3829	TACGGGAGTTTCGATGATGAGAGTGCSCCTCAACCTCCCTTACATCGAACAGGGA	3888
Db	3841	TACGGGAGTTTCGATGATGAGAGTGCSCCTCAACCTCCCTTACATCGAACAGGGA	3900
QY	3889	ATGCAGCTCGCCGAAACAAATTCAAACAGAGCAATCGGTTGCTGCAACAGCCACCAAG	3948
Db	3901	ATGCAGCTCGCCGAAACAAATTCAAACAGAGCAATCGGTTGCTGCAACAGCCACCAAG	3960
QY	3949	CAACGGAGGTGTGCTCCGTTGGTGAATPCCAAAGTGGGACCCCTCGAAGCCTTCTGG	4008
Db	3961	CAACGGAGGTGTGCTCCGTTGGTGAATPCCAAAGTGGGACCCCTCGAAGCCTTCTGG	4020
QY	4009	GCGAAGCATATGTGGAATTTTCATCAGCGGGATCAATATTTTAGCAGGCTTTCACACTG	4068
Db	4021	GCGAAGCATATGTGGAATTTTCATCAGCGGGATCAATATTTTAGCAGGCTTTCACACTG	4080
QY	4069	CTTGGCAACCCCGGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCCGCTC	4128
Db	4081	CCTGGCAACCCCGGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCCGCTC	4140
QY	4129	ACCAACCAATACCTCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGCT	4188
Db	4141	ACCAACCAATACCTCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGCT	4200
QY	4189	CCTCCAGCGTGTCTGCTTTCTGTTCTAGGGCCGGCATCGCTGAGAGGCTGTGGCAGC	4248
Db	4201	CCTCCAGCGTGTCTGCTTTCTGTTCTAGGGCCGGCATCGCTGAGAGGCTGTGGCAGC	4260

QY	4249	ATAGCCCTTGGAAAGTCTGTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGC	4308
Db	4261	ATAGCCCTTGGAAAGTCTGTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGC	4320
QY	4309	CGCCTCGTGGCTTTTAAGGTCATGAGCGCGAGATGCCCTCCACGAGGACCTGTTTAC	4368
Db	4321	CGCCTCGTGGCTTTTAAGGTCATGAGCGCGAGATGCCCTCCACGAGGACCTGTTTAC	4380
QY	4369	CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTCGACGATA	4428
Db	4381	CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTCGACGATA	4440
QY	4429	CTGCGTCGCGACGTGGGCCAGGGAGGGGCTGTGCACTGGATGAAACCGCTGTATGG	4488
Db	4441	CTGCGTCGCGACGTGGGCCAGGGAGGGGCTGTGCACTGGATGAAACCGCTGTATGG	4500
QY	4489	TTGCGTTCCGCGGGTAAACACGTCCTCCCGCCAGCACTATGTGCCTGAGAGCGACCTGCA	4548
Db	4501	TTGCGTTCCGCGGGTAAACACGTCCTCCCGCCAGCACTATGTGCCTGAGAGCGACCTGCA	4560
QY	4549	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCACTACCTCAGCTGCTGAAAGGCTTAC	4608
Db	4561	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCACTACCTCAGCTGCTGAAAGGCTTAC	4620
QY	4609	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGTCTTGG	4668
Db	4621	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGTCTTGG	4680
QY	4669	GATTGGATATGACCGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTGCGG	4728
Db	4681	GATTGGATATGACCGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTGCGG	4740
QY	4729	CGATTGCGGAGTCCCTCTTCTCATGTCAAACGTGGGTACAGGGAGTCTGGCGGGC	4788
Db	4741	CGATTGCGGAGTCCCTCTTCTCATGTCAAACGTGGGTACAGGGAGTCTGGCGGGC	4800
QY	4789	GACGGCATCATGCAAAACCACTGCCCATGTGGAGCAGACATCACGGACATGTGAAAC	4848
Db	4801	GACGGCATCATGCAAAACCACTGCCCATGTGGAGCAGACATCACGGACATGTGAAAC	4860
QY	4849	GGTTCCATGAGGATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCC	4908
Db	4861	GGTTCCATGAGGATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCC	4920
QY	4909	ATTAAACGGTACACACGGGCTTGCACGCCCTTCCCGGGCCAAATATTCTAGGGCG	4968
Db	4921	ATTAAACGGTACACACGGGCTTGCACGCCCTTCCCGGGCCAAATATTCTAGGGCG	4980
QY	4969	CTGTGGCGGTGGCTGTGAGGAGTACGTGGAGGTACGGGGTGGGGATTTCCACTAC	5028
Db	4981	CTGTGGCGGTGGCTGTGAGGAGTACGTGGAGGTACGGGGTGGGGATTTCCACTAC	5040
QY	5029	GTGACGGGCTAGCACTGACCAAGTAAAGTGCCTGTGTGAGGTTCCGGCCCCCGAATTC	5088
Db	5041	GTGACGGGCTAGCACTGACCAAGTAAAGTGCCTGTGTGAGGTTCCGGCCCCCGAATTC	5100
QY	5089	TTTACAGAGTGGATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGAACCCCTCTTA	5148
Db	5101	TTTACAGAGTGGATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGAACCCCTCTTA	5160
QY	5149	CGGAGGAGGTTCATTTCTCGGGCTCAATCAATACCTGTTGGTTCAGCTCCCA	5208
Db	5161	CGGAGGAGGTTCATTTCTCGGGCTCAATCAATACCTGTTGGTTCAGCTCCCA	5220
QY	5209	TGGAGCCGAAACGGACGTAGCAGTGTCTCACTTCCATGTCTACCGACCCCTCCCAATT	5268
Db	5221	TGGAGCCGAAACGGACGTAGCAGTGTCTCACTTCCATGTCTACCGACCCCTCCCAATT	5280
QY	5269	ACGGCGAGAGCGGTAGCGTGGCCAGGGATCTCCCGCTCTCTGGCGAGCTCA	5328
Db	5281	ACGGCGAGAGCGGTAGCGTGGCCAGGGATCTCCCGCTCTCTGGCGAGCTCA	5340
QY	5329	TCAGTAGCCAGCTGTCTGGCGCTTCTTGTAAAGCAACATGACATACCGCTCATGACTCC	5388

Db |||||  
5341 TCAGTAGCGAGCTGCTCGCGCTTCCCTTGAGAGCAACATGCACATACCGCTCATGACTCC 5400  
QY CCGAGCGTGAACCTCATCGAGGCCAACCTCTGTGGCGGAGAGATGGCGGGAAACATC 5448  
Db CCGAGCGTGAACCTCATCGAGGCCAACCTCTGTGGCGGAGAGATGGCGGGAAACATC 5460  
QY ACCCGGTGGAGTCAGAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAAAGCG 5508  
Db ACCCGGTGGAGTCAGAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAAAGCG 5520  
QY GAGGAGGATGAGAGGAGTACCGTTCGGCGGAGATCCCTCGGAGGTCCAGGAAATTC 5568  
Db GAGGAGGATGAGAGGAGTACCGTTCGGCGGAGATCCCTCGGAGGTCCAGGAAATTC 5580  
QY CCTGAGCGATGCCCATATGGSCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTGG 5628  
Db CCTGAGCGATGCCCATATGGSCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTGG 5640  
QY AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCCATTCGCCGCTGCCAAGGCC 5688  
Db AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCCATTCGCCGCTGCCAAGGCC 5700  
QY CCTCGATACCACTCCACGGAGGAGGACGGTTGCTCTGTGAGATCTACCGTGTCT 5748  
Db CCTCGATACCACTCCACGGAGGAGGACGGTTGCTCTGTGAGATCTACCGTGTCT 5760  
QY TCTGCTTTGGCGAGCTGCCACAAAGACCTTCGGCAGCTCCGAAATCTGTCGCCGCTCGAC 5808  
Db TCTGCTTTGGCGAGCTGCCACAAAGACCTTCGGCAGCTCCGAAATCTGTCGCCGCTCGAC 5820  
QY AGCGGACGGGAAAGCGCTCTCTGACGAGCCCTCGAGCAGCGGAGCGGAGTCCGAC 5868  
Db AGCGGACGGGAAAGCGCTCTCTGACGAGCCCTCGAGCAGCGGAGCGGAGTCCGAC 5880  
QY GTTAGTCTGACTCTCTCCATGCCCCCTTGAGGGGAGCGCGGGATCCCGATCTCAGC 5928  
Db GTTAGTCTGACTCTCTCCATGCCCCCTTGAGGGGAGCGCGGGATCCCGATCTCAGC 5940  
QY GACGGTCTTGTCTACCGTAAGCAGGAGGCTAGTAGGAGCTGCTGTGCTCGATG 5988  
Db GACGGTCTTGTCTACCGTAAGCAGGAGGCTAGTAGGAGCTGCTGTGCTCGATG 6000  
QY TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGGCC 6048  
Db TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGGCC 6060  
QY ATCAATGCACTGAGCAACTCTTTGCTCCGTCCACCAACTTGGTCTATGCTACACATCT 6108  
Db ATCAATGCACTGAGCAACTCTTTGCTCCGTCCACCAACTTGGTCTATGCTACACATCT 6120  
QY CGCAGCGCAAGCTCGCGCAGAGAGGTACCTTTGACAGACTGCGAGTCTCGGACGAC 6168  
Db CGCAGCGCAAGCTCGCGCAGAGAGGTACCTTTGACAGACTGCGAGTCTCGGACGAC 6180  
QY CACTACCGGAGCTGCTAAGGAGATGAAGGAGGCGTCCACAGTTAAGGCTAACTT 6228  
Db CACTACCGGAGCTGCTAAGGAGATGAAGGAGGCGTCCACAGTTAAGGCTAACTT 6240  
QY CTATCCGTGGAGAGCTGTAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6288  
Db CTATCCGTGGAGAGCTGTAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6300  
QY TATGGGCAAAAGACCTCGGAACTTATCAGCAAGGCGGTAAACACATCCGCTCCGTG 6348  
Db TATGGGCAAAAGACCTCGGAACTTATCAGCAAGGCGGTAAACACATCCGCTCCGTG 6360  
QY TGAAGGACTTGTGGAAGACTGAGACCAATTTGACCAACCAATTCATGCGCAAAAT 6408  
Db TGAAGGACTTGTGGAAGACTGAGACCAATTTGACCAACCAATTCATGCGCAAAAT 6420  
QY GAGGTTTTCTGGTCCAAACAGAGAGGGGGCCGACGCTGCCTTATCGTATTC 6468

Db |||||  
6421 GAGGTTTTCTCGTCCAAACAGAGAGGGGGCCCAAGCCAGCTCGCTTATCGTATTC 6480  
QY CCAGATTTGGGGTTTGTGTGTGCGAGAAAAATGGCCCTTTAGATGTGTGTCTCCACCTTC 6528  
Db CCAGATTTGGGGTTTGTGTGTGCGAGAAAAATGGCCCTTTAGATGTGTGTCTCCACCTTC 6540  
QY CCTCAGGCGTGTGCGGCTCTTCATACGATTCGAATCTCTCTGAGCAGCGGTCGAG 6588  
Db CCTCAGGCGTGTGCGGCTCTTCATACGATTCGAATCTCTCTGAGCAGCGGTCGAG 6600  
QY TTCTCTGTGAATGCTCGTGAAGCGAAGAAATGCCCCATATGGGCTTCGCATATGACACCCGC 6648  
Db TTCTCTGTGAATGCTCGTGAAGCGAAGAAATGCCCCATATGGGCTTCGCATATGACACCCGC 6660  
QY TGTTTTGTACTCAACGGTCACTGAGAAATGACATTCGCTGTGTGAGGAGTCAATCTTACCAATGT 6708  
Db TGTTTTGTACTCAACGGTCACTGAGAAATGACATTCGCTGTGTGAGGAGTCAATCTTACCAATGT 6720  
QY TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAAGGTCTCACAGAGCGGCTTTACATC 6768  
Db TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAAGGTCTCACAGAGCGGCTTTACATC 6780  
QY GGGGGCCCCCTGACTTAATTTCTAAAGGGCAGAACTCGCGGCTATCGCGGTCGCCGCGAGC 6828  
Db GGGGGCCCCCTGACTTAATTTCTAAAGGGCAGAACTCGCGGCTATCGCGGTCGCCGCGAGC 6840  
QY GGTGTACTGACGACGAGCTGCGGTAAATCCCTCAATGTCTTGAAGGCGCGCTCGGGCC 6888  
Db GGTGTACTGACGACGAGCTGCGGTAAATCCCTCAATGTCTTGAAGGCGCGCTCGGGCC 6900  
QY TGTGAGCTGCGGAGCTCCAGGACTGACGATGCTGATGCGGAGACGACCTTGTGCTT 6948  
Db TGTGAGCTGCGGAGCTCCAGGACTGACGATGCTGATGCGGAGACGACCTTGTGCTT 6960  
QY ATCTCTGAAAAGCGCGGAGACCCAGAGGACGAGGCGAGCTACGGGCTCTTACCGAGGCT 7008  
Db ATCTCTGAAAAGCGCGGAGACCCAGAGGACGAGGCGAGCTACGGGCTCTTACCGAGGCT 7020  
QY ATGACTAGATCTCTGCCCCCTCGGGACCCGCCAAACCAAGATACGACTTGGAGTTG 7068  
Db ATGACTAGATCTCTGCCCCCTCGGGACCCGCCAAACCAAGATACGACTTGGAGTTG 7080  
QY ATAAATCATGCTCTCCAAATGTGTGCTGCGCAGCATGCTTGGCAAAAAGGTTAC 7128  
Db ATAAATCATGCTCTCCAAATGTGTGCTGCGCAGCATGCTTGGCAAAAAGGTTAC 7140  
QY TATCTCACCCGTGACCCCAACCCCTTTCGCGGGCTGCGTGGGAGACAGTACAC 7188  
Db TATCTCACCCGTGACCCCAACCCCTTTCGCGGGCTGCGTGGGAGACAGTACAC 7200  
QY ACTCAGTCAATTCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGGCAAG 7248  
Db ACTCAGTCAATTCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGGCAAG 7260  
QY ATGATCTGATGACTCAATTTCTTCCATCTTCTCCATCTTCTAGCTCAGAAACCACTTGA 7308  
Db ATGATCTGATGACTCAATTTCTTCCATCTTCTCCATCTTCTAGCTCAGAAACCACTTGA 7320  
QY CTAGATTTGTGACTTACGCGGCTGTTACTCTCCTTGAAGCCACTTGACCTACAGATC 7368  
Db CTAGATTTGTGACTTACGCGGCTGTTACTCTCCTTGAAGCCACTTGACCTACAGATC 7380  
QY ATTCAACGACTCCATGGCTTACGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7428  
Db ATTCAACGACTCCATGGCTTACGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7440  
QY AATAGGCTGCTTCAATGCTCAGGAACTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7488  
Db AATAGGCTGCTTCAATGCTCAGGAACTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7500  
QY CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGCG 7548  
Db CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGCG 7560



1309 TCACCAATATTGCGCTCTTTTGGCAATGTGAGGGCCGGAAACCTGGCCCTGTCTTCTTG 1368  
1321 TCACCAATATTGCGCTCTTTTGGCAATGTGAGGGCCGGAAACCTGGCCCTGTCTTCTTG 1380  
1369 ACGAGCATTCCTAGGGGTCTTCCCTCTCGCAAGGAAATCAAGGTCTGTTGAATGTC 1428  
1381 ACGAGCATTCCTAGGGGTCTTCCCTCTCGCAAGGAAATCAAGGTCTGTTGAATGTC 1440  
1429 GTGAAGGAGCAGTTCCTCTGGAAGTCTTCTGAAGCAAAACACGTCGTAGCGACCTT 1488  
1441 GTGAAGGAGCAGTTCCTCTGGAAGTCTTCTGAAGCAAAACACGTCGTAGCGACCTT 1500  
1489 TCGAGCGAGCGAAACCCCCCACTGCGGACAGGTGCTCTCGGCGCAAAAGCAAGTGTGA 1548  
1501 TCGAGCGAGCGAAACCCCCCACTGCGGACAGGTGCTCTCGGCGCAAAAGCAAGTGTGA 1560  
1549 TAAAGATACCTGCAAAAGCGGCAAAACCCCACTGCGGACAGGTGCTGAGTGTGATGTTG 1608  
1561 TAAAGATACCTGCAAAAGCGGCAAAACCCCACTGCGGACAGGTGCTGAGTGTGATGTTG 1620  
1609 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1668  
1621 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1680  
1669 GTACCCCATTTGATGGGATCTGATCTGGGGGCTCGGTGACATGCTTTACATGTTTATG 1728  
1681 GTACCCCATTTGATGGGATCTGATCTGGGGGCTCGGTGACATGCTTTACATGTTTATG 1740  
1729 TCGAGGTTAAAAAGCTCTAGGCCCCCGAAACCAAGGGGACGTTGTTTCTTTGAAAAA 1788  
1741 TCGAGGTTAAAAAGCTCTAGGCCCCCGAAACCAAGGGGACGTTGTTTCTTTGAAAAA 1800  
1789 CACGATAATACCATGCGGCTATTACGGCTACTCCCAACAGACGCGAGGCTTACTTGGC 1848  
1801 CACGATAATACCATGCGGCTATTACGGCTACTCCCAACAGACGCGAGGCTTACTTGGC 1860  
1849 TGCATCATCACTAGCTACAGCCCGGACAGGAACAGGTGAGGGGAGGTCCAAAGTG 1908  
1861 TGCATCATCACTAGCTACAGCCCGGACAGGAACAGGTGAGGGGAGGTCCAAAGTG 1920  
1909 GTCTCCACCGCAACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTTGGACTGTC 1968  
1921 GTCTCCACCGCAACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTTGGACTGTC 1980  
1969 TATCATGTGTCGGGTCAAAAGACCTTTGCGGCCCAAGGGCCCAATCAACCCAAATGTAC 2028  
1981 TATCATGTGTCGGGTCAAAAGACCTTTGCGGCCCAAGGGCCCAATCAACCCAAATGTAC 2040  
2029 ACCAATGTGACCAAGACCTCTGTCGGCTGGCAAGCGCCCCCGGGCGGTCTCTTGACA 2088  
2041 ACCAATGTGACCAAGACCTCTGTCGGCTGGCAAGCGCCCCCGGGCGGTCTCTTGACA 2100  
2089 CCATGACCTGCGGACGTCGACCTTTACTTGGTCAAGAGGATGCGGATGTCATTCCG 2148  
2101 CCATGACCTGCGGACGTCGACCTTTACTTGGTCAAGAGGATGCGGATGTCATTCCG 2160  
2149 GTGCGCGGCGGGGACAGCAGGGGAGCTTACTCTCCCGACGGCCGCTCTCTACTTG 2208  
2161 GTGCGCGGCGGGGACAGCAGGGGAGCTTACTCTCCCGACGGCCGCTCTCTACTTG 2220  
2209 AAGGGCTCTTCGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTGAGGATCTTTCGG 2268  
2221 AAGGGCTCTTCGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTGAGGATCTTTCGG 2280  
2269 GCTGCGGTGTGACCCGAGGGGTTGGAGCGGTGGACTTTGTACCGTCGAGTCTATG 2328  
2281 GCTGCGGTGTGACCCGAGGGGTTGGAGCGGTGGACTTTGTACCCGTCGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCGGTCTTTCAGCGCAAACTCTGTCCTCGGCGGTACCGGAG 2388  
2341 GAAACCACTATGCGGTCCCGGTCTTTCAGCGCAAACTCTGTCCTCGGCGGTACCGGAG 2400  
2389 ACATTCAGGTGCGCCATCTACAGGCCCTTACTGTTAGCGGCAAGACATAAGGTGCGG 2448

2401 ACATTCAGGTGCGCCATCTACGCCCTTACTGTTAGCGGCAAGACATAAGTGGCG 2460  
2449 GCTGCGTATGCAAGCCCAAGGGTATAAGGTGTTCTTCTGAAACCCGTCGTCGCGCCACC 2508  
2461 GCTGCGTATGCAAGCCCAAGGGTATAAGGTGTTCTTCTGAAACCCGTCGTCGCGCCACC 2520  
2509 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGTTATGACCCCTAAACATGAAACCGG 2568  
2521 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGTTATGACCCCTAAACATGAAATCGG 2580  
2569 GTAAAGCAACATCACACGGTCCGCCCATCAGCTACTCCACCTATGGCAAGTTCTTGGC 2628  
2581 GTAAAGCAACATCACACGGTCCGCCCATCAGCTACTCCACCTATGGCAAGTTCTTGGC 2640  
2629 GACGGTGGTGTCTCTGGGGGCGCTATGACATCATATAATATGTGATGAGTGCACCTCACT 2688  
2641 GACGGTGGTGTCTCTGGGGGCGCTATGACATCATATAATATGTGATGAGTGCACCTCACT 2700  
2689 GACTCGACCACTATCTCTGGGCACTCGGCACAGTCTTGGAACAAGCGGAGCGGTGGAGCG 2748  
2701 GACTCGACCACTATCTCTGGGCACTCGGCACAGTCTTGGAACAAGCGGAGCGGTGGAGCG 2760  
2749 CGACTCGTCTGCTCGGCACCGCTAGCCTCGGGATCGGTCAACCTGCGCACATCCAAAC 2808  
2761 CGACTCGTCTGCTCGGCACCGCTAGCCTCGGGATCGGTCAACCTGCGCACATCCAAAC 2820  
2809 ATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCGAAAGCAATCCCC 2868  
2821 ATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCGAAAGCAATCCCC 2880  
2869 ATCGAGACCATCAAGGGGGAGGACCTCTTTTCTGCGCANTTCAAGAAATGTGAT 2928  
2881 ATCGAGACCATCAAGGGGGAGGACCTCTTTTCTGCGCANTTCAAGAAATGTGAT 2940  
2929 GAGCTCGCGCGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 2988  
2941 GAGCTCGCGCGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 3000  
2989 GATGTATCGTCTATACCACTAGCGGAGAGCTATTTGCTAGCAACGAGCGCTTAATG 3048  
3001 GATGTATCGTCTATACCACTAGCGGAGAGCTATTTGCTAGCAACGAGCGCTTAATG 3060  
3049 ACGGGCTTTACCGGCGATTTGACTCAGTGTGATGCTGAATACATGTCACCCAGACA 3108  
3061 ACGGGCTTTACCGGCGATTTGACTCAGTGTGATGCTGAATACATGTCACCCAGACA 3120  
3109 GTGCACTTCAGCTCGGACCGGCTTACCATTGAGACGACCGCTGCGCAAGACGCG 3168  
3121 GTGCACTTCAGCTCGGACCGGCTTACCATTGAGACGACCGCTGCGCAAGACGCG 3180  
3169 GTGCACTTCAGCTCGGACCGGCGAGGAGTGTGAGGGGAGGATGGGCAATTTACAGTTT 3228  
3181 GTGCACTTCAGCTCGGACCGGCGAGGAGTGTGAGGGGAGGATGGGCAATTTACAGTTT 3240  
3229 GTGACTCCAGGAGAACGGGCTTCGGGCAATGTCGATTCCTCGTTCCTGTCGAGTGTAT 3288  
3241 GTGACTCCAGGAGAACGGGCTTCGGGCAATGTCGATTCCTCGTTCCTGTCGAGTGTAT 3300  
3289 GAGCGGGCTGTGTTGTGACGAGCTCACGCCCGGAGACCTCAGTTAGTTTGGGGCT 3348  
3301 GAGCGGGCTGTGTTGTGACGAGCTCACGCCCGGAGACCTCAGTTAGTTTGGGGCT 3360  
3349 TACTTAAACACACCGGGTTCGGCTTCGCGAGGACCATCTGAGTGTCTGGGAGACGCTC 3408  
3361 TACTTAAACACACCGGGTTCGGCTTCGCGAGGACCATCTGAGTGTCTGGGAGACGCTC 3420  
3409 TTTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGAC 3468  
3421 TTTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGAC 3480  
3469 AACTTCCCTTACTGTTAGCATACAGGCTACGGTGTGCGCGAGGCTCAGGCTCCACCT 3528



Db 3481 AACCTCCCTACTCTGTTAGCATACACAGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCT 3540  
Qy 3529 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCCTACCGTGCACGGGCA 3588  
Db 3541 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCCTACCGTGCACGGGCA 3600  
Qy 3589 AGCCGCCCTGCTGTATAGGCTGGGAGCCGTTCAAAAACGAGGTTACTACACACACACCCCAT 3648  
Db 3601 AGCCGCCCTGCTGTATAGGCTGGGAGCCGTTCAAAACGAGGTTACTACACACACACCCCAT 3660  
Qy 3649 ACCAAATACATCATGGCATGCAATGTGCGGTGACCTGGAGGTGCTCAACGACACCTGGGTG 3708  
Db 3661 ACCAAATACATCATGGCATGCAATGTGCGGTGACCTGGAGGTGCTCAACGACACCTGGGTG 3720  
Qy 3709 CTGTTAGCGGAGTCCCTAGCAGCTCTGGCCGGTATTCCTGACACACAGGCGAGCTGGTC 3768  
Db 3721 CTGTTAGCGGAGTCCCTAGCAGCTCTGGCCGGTATTCCTGACACACAGGCGAGCTGGTC 3780  
Qy 3769 ATTGTGGCAGGATCATCTTGTCCGAAAGCCGGCCCATCATTCCTCGACAGGGAAGTCCCTT 3828  
Db 3781 ATTGTGGCAGGATCATCTTGTCCGAAAGCCGGCCCATCATTCCTCGACAGGGAAGTCCCTT 3840  
Qy 3829 TACCGGAGTTCGATGAGATGAAGAGTGGGCTCACACCTCCCTTAATCGAAACAGGGA 3888  
Db 3841 TACCGGAGTTCGATGAGATGAAGAGTGGGCTCACACCTCCCTTAATCGAAACAGGGA 3900  
Qy 3889 ATGACAGCTCGCGCAACATTCAAAACAGAGGCAATCGGTTGCTGCAACACAGCCCAAG 3948  
Db 3901 ATGACAGCTCGCGCAACATTCAAAACAGAGGCAATCGGTTGCTGCAACACAGCCCAAG 3960  
Qy 3949 CAAGCGGAGGTGCTGCTCCGTTGGTGAATTCAAAGTGGCGGACCCCTCGAAGCCTTCGG 4008  
Db 3961 CAAGCGGAGGTGCTGCTCCGTTGGTGAATTCAAAGTGGCGGACCCATCGAAGCCTTCGG 4020  
Qy 4009 GCGAAGCATATGTGGAATTCATCAGCGGATACAAATATTAGCAGGCTTGTCCACTGTG 4068  
Db 4021 GCGAAGCATATGTGGAATTCATCAGCGGATACAAATATTAGCAGGCTTGTCCACTGTG 4080  
Qy 4069 CCTGGCAACCCCGCATAGCATCACTGATGCAATTCACAGCTCTATCACAGCCCGCTC 4128  
Db 4081 CCTGGCAACCCCGCATAGCATCACTGATGCAATTCACAGCTCTATCACAGCCCGCTC 4140  
Qy 4129 ACCACCCCAACATACCTCCTGTTAAATCTCTGGGGGATGGGTGGCCGCCCAACTGCT 4188  
Db 4141 ACCACCCCAACATACCTCCTGTTAAATCTCTGGGGGATGGGTGGCCGCCCAACTGCT 4200  
Qy 4189 CCTCCAGCGCTGCTTCTGCTTCTAGCGCGCGCATCGCTGAGCGGCTGTGGCAGC 4248  
Db 4201 CCTCCAGCGCTGCTTCTGCTTCTAGCGCGCGCATCGCTGAGCGGCTGTGGCAGC 4260  
Qy 4249 ATAGGCTTGGGAAGTCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGCGAGGC 4308  
Db 4261 ATAGGCTTGGGAAGTCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGCGAGGC 4320  
Qy 4309 GCGTCTGGCTTTAAGGTATGAGCGCGCAGATGCCCTCCACCGAGACCTGGTTAAC 4368  
Db 4321 GCGTCTGGCTTTAAGGTATGAGCGCGCAGATGCCCTCCACCGAGACCTGGTTAAC 4380  
Qy 4369 CTACTCCCTGTATCTCTCCCTGGCCCTAGTCTGCGGGTCTGTGGCAGCGATA 4428  
Db 4381 CTACTCCCTGTATCTCTCCCTGGCCCTAGTCTGCGGGTCTGTGGCAGCGATA 4440  
Qy 4429 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTGCAAGTGGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTGCAAGTGGATGAACCGGCTGATAGCG 4500  
Qy 4489 TTCCCTTCGCGGGTAAACCACTCTCCCGCAGCACTATGTCCTGAGAGCGACGCTGCA 4548  
Db 4501 TTCCCTTCGCGGGTAAACCACTCTCCCGCAGCACTATGTCCTGAGAGCGACGCTGCA 4560  
Qy 4549 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTCTGAGAGGCTTCAC 4608  
Db 4561 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTCTGAGAGGCTTCAC 4620

Qy 4609 CAGTGGATCAACAGGAGCTGCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACAGGAGCTGCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTGG 4680  
Qy 4669 GATTGGATATGACAGGCTGTGACTGATTTCAAGACTGCTGCCTCAGTCCAAAGTCTCTGCCG 4728  
Db 4681 GATTGGATATGACAGGCTGTGACTGATTTCAAGACTGCTGCCTCAGTCCAAAGTCTCTGCCG 4740  
Qy 4729 CGATTGCGGGAGTCCCTTCTTCTCATGTCAACCTGGGTACAAGGGAGTCTGGCGGGC 4788  
Db 4741 CGATTGCGGGAGTCCCTTCTTCTCATGTCAACCTGGGTACAAGGGAGTCTGGCGGGC 4800  
Qy 4789 GACGCGCATCATCAAAACCACTGCCATGTGGAGCAGACATCACCGGATGTGAAAAAC 4848  
Db 4801 GACGCGCATCATCAAAACCACTGCCATGTGGGGCACAGATCACCGGATGTGAAAAAC 4860  
Qy 4849 GGTTCATGAGGATCGTGGGGCTAGGACCTGTAGTAACACGTGCATGGAACTTCCCC 4908  
Db 4861 GGTTCATGAGGATCGTGGGGCTAGGACCTGTAGTAACACGTGCATGGAACTTCCCC 4920  
Qy 4909 ATTAAACGGTACACACCGGCGCTGTGACGCCCTCCCGGGCCAAAATTATTTAGGGC 4968  
Db 4921 ATTAAACGGTACACACCGGCGCTGTGACGCCCTCCCGGGCCAAAATTATTTAGGGC 4980  
Qy 4969 CTGTGGCGGGTGGCTGTGAGAGTACGTGAGGTTACCGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGGTGGCTGTGAGAGTACGTGAGGTTACCGGGTGGGGATTTCCACTAC 5040  
Qy 5029 GTGACGGGCATGACACCACTGACAAATAAAGTGCCTGTACAGTTCCGGGCCCCAAATTC 5088  
Db 5041 GTGACGGGCATGACACCACTGACAAATAAAGTGCCTGTACAGTTCCGGGCCCCAAATTC 5100  
Qy 5089 TTCTCAGAAAGTGGATGGGGTGGTTCACAGGTACGCTCCAGGTGCAAAACCCCTCTTA 5148  
Db 5101 TTCTCAGAAAGTGGATGGGGTGGTTCACAGGTACGCTCCAGGTGCAAAACCCCTCTTA 5160  
Qy 5149 CGGAGGAGGTGACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCAGCTCCCA 5208  
Db 5161 CGGAGGAGGTGACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCAGCTCCCA 5220  
Qy 5209 TCGAGGCCGCAACCGGAGTAGCAGTGTCACTTCCATGTCTCAGCGACCCCTCCACAT 5268  
Db 5221 TCGAGGCCGCAACCGGAGTAGCAGTGTCACTTCCATGTCTCAGCGACCCCTCCACAT 5280  
Qy 5269 ACGGCGGAGAGCGCTAAGCTAGGCTGGCGAGGGATCTCCCGCTCTTGGCGAGCTCA 5328  
Db 5281 ACGGCGGAGAGCGCTAAGCTAGGCTGGCGAGGGATCTCCCGCTCTTGGCGAGCTCA 5340  
Qy 5329 TCAGCTAGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGACATACCGCTCATGACTCC 5388  
Db 5341 TCAGCTAGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGACATACCGCTCATGACTCC 5400  
Qy 5389 CCGAGCTGACCTCATCTGAGGCCCAACCTCTGTGGGGCAGGAGATGGGCGGAAATC 5448  
Db 5401 CCGAGCTGACCTCATCTGAGGCCCAACCTCTGTGGGGCAGGAGATGGGCGGAAATC 5460  
Qy 5449 ACCCGGTGAGTGCAGAAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAGCG 5508  
Db 5461 ACCCGGTGAGTGCAGAAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAGCG 5520  
Qy 5509 GAGGAGATGAGAGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGGAAATC 5568  
Db 5521 GAGGAGATGAGAGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGGAAATC 5580  
Qy 5569 CTTGAGCGATGCCATATATGGGACCGCCGGATTAACCCCTCCACTGTTTAGAGTCTCTGG 5628  
Db 5581 CTTGAGCGATGCCATATATGGGACCGCCGGATTAACCCCTCCACTGTTTAGAGTCTCTGG 5640  
Qy 5629 AAGGACCGGACTACGCTCCCTCCAGTGTACAGGGTGTCCATTGCGGCTCCCAAGGCC 5688  
Db 5641 AAGGACCGGACTACGCTCCCTCCAGTGTACAGGGTGTCCATTGCGGCTCCCAAGGCC 5700



QY 5689 CCTCCGATACACCTCCAGGAGGAGGAGCGTGTCTCTGTCTCAGAAATCTTACCGTGTCT 5748  
Db |||||  
QY 5701 CCTCGATACACCTTCCAGGAGGAGGAGCGTGTCTCTGTCTCAGAAATCTTACCGTGTCT 5760  
Db |||||  
QY 5749 TCTGCTTTGGGGAGCTCGCCAAAGACCTTCGCGCAGCTCCGAATCTGTCGCGCGTTCGAC 5808  
Db |||||  
QY 5761 TCTGCTTTGGGGAGCTCGCCAAAGACCTTCGCGCAGCTCCGAATCTGTCGCGCGTTCGAC 5820  
Db |||||  
QY 5809 AGCGCAGCGGAAACCGCTCTCTCTGACCAAGCCCTCCGACGAGCGGAGACCGCGGATCCGAC 5868  
Db |||||  
QY 5821 AGCGCAGCGGAAACCGCTCTCTCTGACCAAGCCCTCCGACGAGCGGAGACCGCGGATCCGAC 5880  
Db |||||  
QY 5869 GTTGAGTGTACTCTCTCATGCGCCCTCTTGAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db |||||  
QY 5881 GTTGAGTGTACTCTCTCATGCGCCCTCTTGAGGGGAGCGGGGATCCCGATCTCAGC 5940  
Db |||||  
QY 5929 GACGGGTCTTGTGTCTTACCGTAAGCAGGAGGCTAGTGAGGAGCGTCTGTCTGTCTCGATG 5988  
Db |||||  
QY 5941 GACGGGTCTTGTGTCTTACCGTAAGCAGGAGGCTAGTGAGGAGCGTCTGTCTGTCTCGATG 6000  
Db |||||  
QY 5989 TCCTACATGACAGCGGCGCTCTCTCTGACCAAGCCCTCCGACGAGCGGAGACCGCGGATCCGAC 6048  
Db |||||  
QY 6001 TCCTACATGACAGCGGCGCTCTCTCTGACCAAGCCCTCCGACGAGCGGAGACCGCGGATCCGAC 6060  
Db |||||  
QY 6049 ATCAATGCACTGAGCAACTCTTTGCTCGCTCAGCAAACTTGTGTCTATGCTACCAATCT 6108  
Db |||||  
QY 6061 ATCAATGCACTGAGCAACTCTTTGCTCGCTCAGCAAACTTGTGTCTATGCTACCAATCT 6120  
Db |||||  
QY 6109 CGCAGCGCAAGCTCTCGGCGAGAAAGGTCACTTTGACAGACTGCGAGTCTCTGAGCAGC 6168  
Db |||||  
QY 6121 CGCAGCGCAAGCTCTCGGCGAGAAAGGTCACTTTGACAGACTGCGAGTCTCTGAGCAGC 6180  
Db |||||  
QY 6169 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGGTCCACAGTTAAGGCTAAACTT 6228  
Db |||||  
QY 6181 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGGTCCACAGTTAAGGCTAAACTT 6240  
Db |||||  
QY 6229 CTATCGGTGAGGAGCGTCTCAAGGAGATGAAGGCGAAGCGGTCCACAGTTAAGGCTAAACTT 6288  
Db |||||  
QY 6241 CTATCGGTGAGGAGCGTCTCAAGGAGATGAAGGCGAAGCGGTCCACAGTTAAGGCTAAACTT 6300  
Db |||||  
QY 6289 TATGGGCAAGAGCTCTCGGAACTTATCCAGCAAGCGGTAAACCACTCCGCTCCGCTG 6348  
Db |||||  
QY 6301 TATGGGCAAGAGCTCTCGGAACTTATCCAGCAAGCGGTAAACCACTCCGCTCCGCTG 6360  
Db |||||  
QY 6349 TGGAGGACTTCTGCGAAGCACTGAGACACCAATTTGACACCACTATGCGCAAAAT 6408  
Db |||||  
QY 6361 TGGAGGACTTCTGCGAAGCACTGAGACACCAATTTGACACCACTATGCGCAAAAT 6420  
Db |||||  
QY 6409 GAGGTTTCTGCTGCTCAACGAGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6468  
Db |||||  
QY 6421 GAGGTTTCTGCTGCTCAACGAGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6480  
Db |||||  
QY 6469 CCAGATTTGGGGTCTGCTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCCCTC 6528  
Db |||||  
QY 6481 CCAGATTTGGGGTCTGCTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCCCTC 6540  
Db |||||  
QY 6529 CCTCAGCGCGTGTGGGTCTTCCATCGGATTCATATCTCTCCTCGACAGCGGGTTCGAG 6588  
Db |||||  
QY 6541 CCTCAGCGCGTGTGGGTCTTCCATCGGATTCATATCTCTCCTCGACAGCGGGTTCGAG 6600  
Db |||||  
QY 6589 TTCTCGTGAATCCCTGGAAGCGAAGATGCCCTATGGCTTCGCATATGACACCCGC 6648  
Db |||||  
QY 6601 TTCTCGTGAATCCCTGGAAGCGAAGATGCCCTATGGCTTCGCATATGACACCCGC 6660  
Db |||||  
QY 6649 TGTGTTTGACTCAACGGTCACTGAGATGACATCCGTGTGTGAGGAGTCAATCTACCAATGT 6708  
Db |||||  
QY 6661 TGTGTTTGACTCAACGGTCACTGAGATGACATCCGTGTGTGAGGAGTCAATCTACCAATGT 6720  
Db |||||  
QY 6709 TGTGACTTGGCCCCCGAAGCAGACAGGCGCATTAAGTTCGCTCAGAGCGGGCTTTACATC 6768  
Db |||||  
QY 6721 TGTGACTTGGCCCCCGAAGCAGACAGGCGCATTAAGTTCGCTCAGAGCGGGCTTTACATC 6780  
Db |||||  
QY 6769 GGGGGCCCCCTGACTAAATCTTAAGGCGAGAACTGCGGCTATCGCGGTGCCCGCGAGC 6828  
Db |||||

Db 6781 GGGGGCCCCCTGACTAAATCTTAAGGCGAGAACTGCGGCTATCGCGGTGCCCGGAGC 6840  
QY |||||  
QY 6829 GGTGTACTGACGACACAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCCGCTCGGGCC 6888  
Db |||||  
QY 6841 GGTGTACTGACGACACAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCCGCTCGGGCC 6900  
Db |||||  
QY 6889 TGTGAGCTGCGAAGCTCGAGACTGCGAGTGTCTGATGCGGAGACGACCTTGTCTGT 6948  
Db |||||  
QY 6901 TGTGAGCTGCGAAGCTCGAGACTGCGAGTGTCTGATGCGGAGACGACCTTGTCTGT 6960  
Db |||||  
QY 6949 ATCTGTGAAAGCGCGGAGACCAAGAGGAGGAGGAGCTTACGGGCTTCCAGGAGGCT 7008  
Db |||||  
QY 6961 ATCTGTGAAAGCGCGGAGACCAAGAGGAGGAGGAGCTTACGGGCTTCCAGGAGGCT 7020  
Db |||||  
QY 7009 ATGACTAGATACTCTGCCCCCTCGGGACCCGCGCAAAACCAAGATAAGCTTGGAGTTG 7068  
Db |||||  
QY 7021 ATGACTAGATACTCTGCCCCCTCGGGACCCGCGCAAAACCAAGATAAGCTTGGAGTTG 7080  
Db |||||  
QY 7069 ATAAATCATGCTCTCTCCAAATGTGTCTGCGCGCAGATGATCTGGCAAAAGGTTGATC 7128  
Db |||||  
QY 7081 ATAAATCATGCTCTCTCCAAATGTGTCTGCGCGCAGATGATCTGGCAAAAGGTTGATC 7140  
Db |||||  
QY 7129 TATCTCACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGAGAGCAGCTAGACAC 7188  
Db |||||  
QY 7141 TATCTCACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGAGAGCAGCTAGACAC 7200  
Db |||||  
QY 7189 ACTCCAGTCAATCTCTGCTAGGCAACATCATATATGCGGCCACCTTGTGGGCAAG 7248  
Db |||||  
QY 7201 ACTCCAGTCAATCTCTGCTAGGCAACATCATATATGCGGCCACCTTGTGGGCAAG 7260  
Db |||||  
QY 7249 ATGATCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAANAAGC 7308  
Db |||||  
QY 7261 ATGATCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAANAAGC 7320  
Db |||||  
QY 7309 CTAGATTTGACATCTACGGGCGCTGTTACTCCATTTAGCCACTTACCTCAGATC 7368  
Db |||||  
QY 7321 CTAGATTTGACATCTACGGGCGCTGTTACTCCATTTAGCCACTTACCTCAGATC 7380  
Db |||||  
QY 7369 ATTTCAACGATCTCATGGCTTAGCGCATTTTCACTCCATGATTTCTCCAGTGTAGATC 7428  
Db |||||  
QY 7381 ATTTCAACGATCTCATGGCTTAGCGCATTTTCACTCCATGATTTCTCCAGTGTAGATC 7440  
Db |||||  
QY 7429 AATAGGTTGCTTATGCTCTAGGAAAATTGGGTACCGCCCTTGGAGTCTGGAGCAT 7488  
Db |||||  
QY 7441 AATAGGTTGCTTATGCTCTAGGAAAATTGGGTACCGCCCTTGGAGTCTGGAGCAT 7500  
Db |||||  
QY 7489 CGGGCCAGAAAGTGTCCGCGCTAGGCTTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGC 7548  
Db |||||  
QY 7501 CGGGCCAGAAAGTGTCCGCGCTAGGCTTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGC 7560  
Db |||||  
QY 7549 AAGTACTCTTCAACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATFCCCGGTGCG 7608  
Db |||||  
QY 7561 AAGTACTCTTCAACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATFCCCGGTGCG 7620  
Db |||||  
QY 7609 TCCAGTTGGATTTATCCAGCTGCTTGTGTTGTTAGCGGGGAGGACATATATC 7668  
Db |||||  
QY 7621 TCCAGTTGGATTTATCCAGCTGCTTGTGTTGTTAGCGGGGAGGACATATATC 7680  
Db |||||  
QY 7669 AGCTGTCTCGTCCCGACCCCGCTGTTTCTGTTGTTGTTCTCTCTCTCTCTCTAGGG 7728  
Db |||||  
QY 7681 AGCTGTCTCGTCCCGACCCCGCTGTTTCTGTTGTTGTTCTCTCTCTCTCTCTAGGG 7740  
Db |||||  
QY 7729 GTAGGCATCTATCTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGG 7788  
Db |||||  
QY 7741 GTAGGCATCTATCTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGG 7800  
Db |||||  
QY 7789 CAAATCTGTTTTTCCCTTT 7848  
Db |||||  
QY 7801 CAAATCTGTTTTTCCCTTT 7860  
Db |||||  
QY 7849 TTTTCTCTCTTTTTTTTTTCTCTTTTTTCTCTTTTTTCTCTTGTGGTCTCAATCTTAGC 7908  
Db |||||



QY 1609 GAAAGAGTCAATGSGTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAG 1668  
DB 1621 GAAAGAGTCAATGSGTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAG 1680  
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1728  
DB 1681 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1740  
QY 1729 TCGAGGTTAAAAACGTCGTAGGCCCCCGAACCAACGCGGACGTTGTTTCTTTGAAAAA 1788  
DB 1741 TCGAGGTTAAAAACGTCGTAGGCCCCCGAACCAACGCGGACGTTGTTTCTTTGAAAAA 1800  
QY 1789 CAGGATATACCATGGCGCTATTAGCGCTACTCTCCACAGACGCGAGCGCTACTTGGC 1848  
DB 1801 CAGGATATACCATGGCGCTATTAGCGCTACTCTCCACAGACGCGAGCGCTACTTGGC 1860  
QY 1849 TGCATCATCACTAGCTACAGCGCGGACAGGAACCAAGTCCAGGGGAGGTCCAAAGTG 1908  
DB 1861 TGCATCATCACTAGCTACAGCGCGGACAGGAACCAAGTCCAGGGGAGGTCCAAAGTG 1920  
QY 1909 GTCTCACGCAACAATCTTTCTGCGACCTGCTGCTCAATGGCGTGTGTTGGAATGTC 1968  
DB 1921 GTCTCACGCAACAATCTTTCTGCGACCTGCTGCTCAATGGCGTGTGTTGGAATGTC 1980  
QY 1969 TATCATGTCGCGGCTCAAGACCTTTCGCGGCCCAAGGGGCCCAATCAACCAATGTAC 2028  
DB 1981 TATCATGTCGCGGCTCAAGACCTTTCGCGGCCCAAGGGGCCCAATCAACCAATGTAC 2040  
QY 2029 ACCAATGTGACACAGGACCTCGTCGGCTGCGAAGCGCCCCCGGGCGGCTTCCTTGACA 2088  
DB 2041 ACCAATGTGACACAGGACCTCGTCGGCTGCGAAGCGCCCCCGGGCGGCTTCCTTGACA 2100  
QY 2089 CCAATGCACTGCGGCACTCGGACCTTTACTTGGTCAAGGCAATCGGATGCAATCCG 2148  
DB 2101 CCAATGCACTGCGGCACTCGGACCTTTACTTGGTCAAGGCAATCGGATGCAATCCG 2160  
QY 2149 GTGCGCGGGGGGACAGAGGGGAGCTACTCTCCCGAGGCGGCTCTCTACTTTCG 2208  
DB 2161 GTGCGCGGGGGGACAGAGGGGAGCTACTCTCCCGAGGCGGCTCTCTACTTTCG 2220  
QY 2209 AAGGGCTCTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTCTGGGCACTTTCGG 2268  
DB 2221 AAGGGCTCTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTCTGGGCACTTTCGG 2280  
QY 2269 GTGCGGCTGTGACCCGAGGGGTTGCGAAGGCGGTGGAATTTGTAACCGTCTGAGTCTATG 2328  
DB 2281 GTGCGGCTGTGACCCGAGGGGTTGCGAAGGCGGTGGAATTTGTAACCGTCTGAGTCTATG 2340  
QY 2329 GAAACCACTATGGTCCCGGTCCTTACGGCAACTCGTCCCTCGGCGCGTACCGCAG 2388  
DB 2341 GAAACCACTATGGTCCCGGTCCTTACGGCAACTCGTCCCTCGGCGCGTACCGCAG 2400  
QY 2389 ACATTCAGGTGCCCCCTACTACAGCCCCCTACTGTTAGCGGCAAGGACCTAAGGTGCCG 2448  
DB 2401 ACATTCAGGTGCCCCCTACTACAGCCCCCTACTGTTAGCGGCAAGGACCTAAGGTGCCG 2460  
QY 2449 GTGCGGCTGTGACCCGCAAGGTTAAGGTGCTTGTCTCTGAAACCGGTCGTGCGCGCAAC 2508  
DB 2461 GTGCGGCTGTGACCCGCAAGGTTAAGGTGCTTGTCTCTGAAACCGGTCGTGCGCGCAAC 2520  
QY 2509 CTAGGTTTCGGGCGGTATATGCTTAGGGCACATGTTATCGACCTTAACATCAAGACCGGG 2568  
DB 2521 CTAGGTTTCGGGCGGTATATGCTTAGGGCACATGTTATCGACCTTAACATCAAGACCGGG 2580  
QY 2569 GTAAGGACCATCACCGGTGCCCCCTACCACTACTTCCACCTATGGCAAGTTTCTTGGC 2628  
DB 2581 GTAAGGACCATCACCGGTGCCCCCTACCACTACTTCCACCTATGGCAAGTTTCTTGGC 2640  
QY 2629 GACGGTGTGCTCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCAACT 2688  
DB 2641 GACGGTGTGCTCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGAACCAAGGGAGACGGCTGGAGG 2748

DB 2701 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGGCTGGAGG 2760  
QY 2749 GCACTCGTCTGCTCGCCACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAAC 2808  
DB 2761 CGACTCGTCTGCTCGCCACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAAC 2820  
QY 2809 ATCGAGGAGGTGGCTCTGTCTAGCACTGAGAAATATCCCTTTTATGGAAGACCATCCCC 2868  
DB 2821 ATCGAGGAGGTGGCTCTGTCTAGCACTGAGAAATATCCCTTTTATGGAAGACCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCATTTCAAGAAAGAAATGTGAT 2928  
DB 2881 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCATTTCAAGAAAGAAATGTGAT 2940  
QY 2929 GAGCTCGCCCGGAAGCTGTCCGGCTCGGACTCAATGTCTGTAGCATATTTACCGGGGCTT 2988  
DB 2941 GAGCTCGCCCGGAAGCTGTCCGGCTCGGACTCAATGTCTGTAGCATATTTACCGGGGCTT 3000  
QY 2989 GATGATCTCGTATACCACTAGCGGAGACGTCATTTCTGTAGCAACGACGCTCTAATG 3048  
DB 3001 GATGATCTCGTATACCACTAGCGGAGACGTCATTTCTGTAGCAACGACGCTCTAATG 3060  
QY 3049 ACGGGCTTTACCGGCAATTTGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGACA 3108  
DB 3061 ACGGGCTTTACCGGCAATTTGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGACA 3120  
QY 3109 GTCGACTTCAGCTCGAACCCGCTTACCAATGAGACGACCGCTGCGCACAAAGCGG 3168  
DB 3121 GTCGACTTCAGCTCGAACCCGCTTACCAATGAGACGACCGCTGCGCACAAAGCGG 3180  
QY 3169 GTGTCACGCTTCGAGCGGCGAGGACGACTGTGTAGGGGAGGATGGGCAATTTACAGTTT 3228  
DB 3181 GTGTCACGCTTCGAGCGGCGAGGACGACTGTGTAGGGGAGGATGGGCAATTTACAGTTT 3240  
QY 3229 GTGACTTCAGAGGAGAACGGGCTTCGGCATGTTCTCGTTCCTGTGCGAGTGCAT 3288  
DB 3241 GTGACTTCAGAGGAGAACGGGCTTCGGCATGTTCTCGTTCCTGTGCGAGTGCAT 3300  
QY 3289 GACGGGCTGTGCTTTGTPACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGT 3348  
DB 3301 GACGGGCTGTGCTTTGTPACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGT 3360  
QY 3349 TACTTAAACACACGAGGTTCGCTGCGAGGACCATCTGGAGTTCTGGGAGACGCTC 3408  
DB 3361 TACTTAAACACACGAGGTTCGCTGCGAGGACCATCTGGAGTTCTGGGAGACGCTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCAGACTAAGCAGGCGAGGAC 3468  
DB 3421 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCAGACTAAGCAGGCGAGGAC 3480  
QY 3469 AACTTCCCCCTACTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCT 3528  
DB 3481 AACTTCCCCCTACTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCT 3540  
QY 3529 CCATGTGGGACCAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCACGGGCA 3588  
DB 3541 CCATGTGGGACCAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCACGGGCA 3600  
QY 3589 AGCCCCCTCTGCTATAGGCTGGGAGCGGTTTAAACGAGGTTACTACCAACACCCCAT 3648  
DB 3601 AGCCCCCTCTGCTATAGGCTGGGAGCGGTTTAAACGAGGTTACTACCAACACCCCAT 3660  
QY 3649 ACCAATACATCATGSCATGCTGCTGAGCTGAGCTCGTCAACGAGCACTCGGGTG 3708  
DB 3661 ACCAATACATCATGSCATGCTGCTGAGCTGAGCTCGTCAACGAGCACTCGGGTG 3720  
QY 3709 CTGTTAGGCGGAGTCTTAGACGCTCTGGCGGCTATTGCTGACAAACGAGCGGTGTC 3768  
DB 3721 CTGTTAGGCGGAGTCTTAGACGCTCTGGCGGCTATTGCTGACAAACGAGCGGTGTC 3780  
QY 3769 ATTGTGGGAGGATCATCTTGTCCGGAAGCGGCCCATCATTTCCGACAGGAAGTCTT 3828



QY 5989 TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAACCAAGCTGCC 6048  
DB 6001 TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAACCAAGCTGCC 6060  
QY 6049 ATCAATGCACTGAGCAACTCTTTGCTCCGTACCACTTGGTCTATCTCAACATCT 6108  
DB 6061 ATCAATGCACTGAGCAACTCTTTGCTCCGTACCACTTGGTCTATGCTACATCT 6120  
QY 6109 CGCAGCGCAAGCCTCGGCGAGAGAGTCACTTTGACAGACTGCAAGTCTCTGACGAC 6168  
DB 6121 CGCAGCGCAAGCCTCGGCGAGAGAGTCACTTTGACAGACTGCAAGTCTCTGACGAC 6180  
QY 6169 CACTACCGGAGCTGCTCAAGAGATGAAGCGAGGCGCTCCACAGTTAAGCTAACTT 6228  
DB 6181 CACTACCGGAGCTGCTCAAGAGATGAAGCGAGGCGCTCCACAGTTAAGCTAACTT 6240  
QY 6229 CTATCGTGGAGAGCCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6288  
DB 6241 CTATCGTGGAGAGCCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAGAGCCTCGGAAACCTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTG 6348  
DB 6301 TATGGGCAAGAGCCTCGGAAACCTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTG 6360  
QY 6349 TGGAGGACTTCTGCGGCAACCCAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTC 6408  
DB 6361 TGGAGGACTTCTGCGGCAACCCAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTC 6420  
QY 6409 GAGGTTTCTGCGTCCAAACAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTC 6468  
DB 6421 GAGGTTTCTGCGTCCAAACAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTC 6480  
QY 6469 CCAGATTTGGGGTTCGTGTGCGAGAAATAGGCCCTTTACATGATGCTCCACCCCTC 6528  
DB 6481 CCAGATTTGGGGTTCGTGTGCGAGAAATAGGCCCTTTACATGATGCTCCACCCCTC 6540  
QY 6529 CCTCAGGCGCTGATGGCTCTCATACGATTCATATCTCTCGAGAGGGGTCGAG 6588  
DB 6541 CCTCAGGCGCTGATGGCTCTCATACGATTCATATCTCTCGAGAGGGGTCGAG 6600  
QY 6589 TTCCTGGTGAATGCCCTGGAAGCGAAGAAATGCCCTTATGGGCTTCGCATATGACACCCGC 6648  
DB 6601 TTCCTGGTGAATGCCCTGGAAGCGAAGAAATGCCCTTATGGGCTTCGCATATGACACCCGC 6660  
QY 6649 TGTTTTGACTCAACCGTCACTAGATGATCATCCGTGTTGAGGAGTCAATCTACCAATGT 6708  
DB 6661 TGTTTTGACTCAACCGTCACTAGATGATCATCCGTGTTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTGGCCCCGAAAGCCAGACAGGCGCATAGGTCGCTCACAGAGCGGCTTTACATC 6768  
DB 6721 TGTGACTTGGCCCCGAAAGCCAGACAGGCGCATAGGTCGCTCACAGAGCGGCTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTAAATTTAAAGGCGCAGAACTGGGGCTATCGCCGGTCCCGCGCGAGC 6828  
DB 6781 GGGGGCCCCCTGACTAAATTTAAAGGCGCAGAACTGGGGCTATCGCCGGTCCCGCGCGAGC 6840  
QY 6829 GGTGTACTGACGCCAGCTGCGGTATACCTCACAATGTTGAGGAGCGGCTCGGCC 6888  
DB 6841 GGTGTACTGACGCCAGCTGCGGTATACCTCACAATGTTGAGGAGCGGCTCGGCC 6900  
QY 6889 TGTGAGCTGCGAGCTCCAGGACTCAGCATGCTCGTATGCGGAGCAGCTTGTGCTT 6948  
DB 6901 TGTGAGCTGCGAGCTCCAGGACTCAGCATGCTCGTATGCGGAGCAGCTTGTGCTT 6960  
QY 6949 ATCTGTGAAGCCGGGAGCCCAAGAGGACGAGGCGGCTACGGGCTTTCACGGAGGCT 7008  
DB 6961 ATCTGTGAAGCCGGGAGCCCAAGAGGACGAGGCGGCTACGGGCTTTCACGGAGGCT 7020  
QY 7009 ATGACTAGATPACTCTGCCCCCTGCGGAGCCCGCCCAACAGAAATGAGCTTGGAGTTG 7068  
DB 7021 ATGACTAGATPACTCTGCCCCCTGCGGAGCCCGCCCAACAGAAATGAGCTTGGAGTTG 7080  
QY 7069 ATAACATCATGCTCTCCCAATGTGTAGTCCGCGCAGCATGCTGGCAAAAGGGTGTAC 7128

DB 7081 ATPAACATCATGCTCTCCATGTGTGTCAGTCCGCGCAGATGCACTGGCAAAAGGGTGTAC 7140  
QY 7129 TATCTACCCCGTGAACCCCAACCCCTTGGCGGGGTGCGTGGAGACAGCTAGACAC 7188  
DB 7141 TATCTACCCCGTGAACCCCAACCCCTTGGCGGGGTGCGTGGAGACAGCTAGACAC 7200  
QY 7189 ACTCAGTCAATCTCGCTAGGCAACATCATCATGATGATGCGCCCACTTGTGGGCAAG 7248  
DB 7201 ACTCAGTCAATCTCGCTAGGCAACATCATCATGATGATGCGCCCACTTGTGGGCAAG 7260  
QY 7249 ATGATCTCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAACTTCAAAAGCC 7308  
DB 7261 ATGATCTCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAACTTCAAAAGCC 7320  
QY 7309 CTAGATTTGATGATCTAGCGGCTCTTACTCCATGAGCCACTTGACCTCAGATC 7368  
DB 7321 CTAGATTTGATGATCTAGCGGCTCTTACTCCATGAGCCACTTGACCTCAGATC 7380  
QY 7369 ATTCAACGACTCATGGCTTAGCGATTTTCACTCCATAGTTACTCTCAGGTGATC 7428  
DB 7381 ATTCAACGACTCATGGCTTAGCGATTTTCACTCCATAGTTACTCTCAGGTGATC 7440  
QY 7429 AATAGGCTGCTTCACTGCTCAGGAACTTGGGTTACCGCCCTTGGAGTCTGGAGACAT 7488  
DB 7441 AATAGGCTGCTTCACTGCTCAGGAACTTGGGTTACCGCCCTTGGAGTCTGGAGACAT 7500  
QY 7489 CGGCGCAGAAAGTGTCCGGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCATTTGGC 7548  
DB 7501 CGGCGCAGAAAGTGTCCGGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCATTTGGC 7560  
QY 7549 AAGTACCTTCAACTGGGCTAGGACCAAGCTCAACTCACTCCAAATCCCGGCTGG 7608  
DB 7561 AAGTACCTTCAACTGGGCTAGGACCAAGCTCAACTCACTCCAAATCCCGGCTGG 7620  
QY 7609 TCCAGTTCGATTTATCCAGCTGGTTCGTGCTGTTACAGGGGGGAGACATATATCAC 7668  
DB 7621 TCCAGTTCGATTTATCCAGCTGGTTCGTGCTGTTACAGGGGGGAGACATATATCAC 7680  
QY 7669 AGCTGTCTCTGCGCGACCCCGCTGTTTCACTGCTGCTACTCTCTCTTTCTGAGG 7728  
DB 7681 AGCTGTCTCTGCGCGACCCCGCTGTTTCACTGCTGCTACTCTCTCTTTCTGAGG 7740  
QY 7729 GTAGGATCTATCTACTCCCAACCGATGAAAGGGGACCTAAACACTCCAGGCCANTAG 7788  
DB 7741 GTAGGATCTATCTACTCCCAACCGATGAAAGGGGACCTAAACACTCCAGGCCANTAG 7800  
QY 7789 CCATCTCTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848  
DB 7801 CCATCTGTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860  
QY 7849 TTTTCTCTTTTTTTTCTTTTTTTTCTTTTTTCTTTTTTCTTTTCTTTTCTTTTCTTTT 7908  
DB 7861 TTTTCTCTTTTTTTTCTTTTTTTTCTTTTTTCTTTTTTCTTTTCTTTTCTTTTCTTTT 7920  
QY 7909 CCTAGTCAGGCTAGCTGTGAAGGTCCTGAGCCGCTTGACTGACAGAGTCTCTATC 7968  
DB 7921 CCTAGTCAGGCTAGCTGTGAAGGTCCTGAGCCGCTTGACTGACAGAGTCTCTATC 7980  
QY 7969 TGGCTCTCTGAGATCAAGT 7989  
DB 7981 TGGCTCTCTGAGATCAAGT 8001

## RESULT 10

AR406045  
LOCUS AR406045 8001 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 16 from patent US 6630343.  
ACCESSION AR406045  
VERSION AR406045.1 GI:40155172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

Unclassified.	
REFERENCE	1 (bases 1 to 8001)
AUTHORS	Bartenschlager,R.
TITLE	Hepatitis C virus culture system
JOURNAL	Patent: US 6630343-A 16 07-OCT-2003;
FEATURES	Location/Qualifiers
source	1..8001
	/organism="unknown"
ORIGIN	/mol_type="genomic DNA"
Query Match 99.5%; Score 7946.2; DB 6; Length 8001;	
Best Local Similarity 99.7%; Pred. No. 0;	
Matches 7976; Conservative 0; Mismatches 13; Indels 12; Gaps 1;	
QY	1 GCCAGCCCCGATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG 60
DB	1 GCCAGCCCCGATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG 60
QY	61 TCTTCAAGCAGAAAGCGTACGATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 120
DB	61 TCTTCAAGCAGAAAGCGTACGATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 120
QY	121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
DB	121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
QY	181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTGGCGGCGCCGCC 240
DB	181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTGGCGGCGCCGCC 240
QY	241 GCGAGACTGTAGCCGAGTGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 300
DB	241 GCGAGACTGTAGCCGAGTGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 300
QY	301 GTGCTTTCGAGTCCCGGAGGCTCTGAGACCGGTGACCATGAGCAGCAAGATCTTAAC 360
DB	301 GTGCTTTCGAGTCCCGGAGGCTCTGAGACCGGTGACCATGAGCAGCAAGATCTTAAC 360
QY	361 CTCGAAGAAACCAAA-----GGGCGGCCATGATTGAACAAGATGGATGC 408
DB	361 CTCGAAGAAACCAAAACGTAACCAACGCGCGCCCATGATTGAACAAGATGGATGC 420
QY	409 ACGCAGGTTCTCCGCGCGTGGTGGAGAGGCTATTGCGGTATGACTGGGCAACAAGA 468
DB	421 ACGCAGGTTCTCCGCGCGTGGTGGAGAGGCTATTGCGGTATGACTGGGCAACAAGA 480
QY	469 CAATCGGCTGCTGTGATCCCGCGTGTTCGCGCTGTGAGCGCAGGGCGCCCGGTTCTTT 528
DB	481 CAATCGGCTGCTGTGATCCCGCGTGTTCGCGCTGTGAGCGCAGGGCGCCCGGTTCTTT 540
QY	529 TTGTCAGACCGGACTGTCCGTCCTGATGAATGAATGACGAGGAGCGCGGCTAT 588
DB	541 TTGTCAGACCGGACTGTCCGTCCTGATGAATGAATGACGAGGAGCGCGGCTAT 600
QY	589 CGTGGCTGGCCACGACCGGCGTTCCTTTCGCGAGTGTGCTGAGCTGTGCTCAAGCGG 648
DB	601 CGTGGCTGGCCACGACCGGCGTTCCTTTCGCGAGTGTGCTGAGCTGTGCTCAAGCGG 660
QY	649 GAAGGACTGGCTGCTATTGGCGAAGTGGCGGGCAGGATCTCCTGTCTACCTTGCCTTG 708
DB	661 GAAGGACTGGCTGCTATTGGCGAAGTGGCGGGCAGGATCTCCTGTCTACCTTGCCTTG 720
QY	709 CTCCTCCGGAAGATGATCCATATGCTGATGATGAATGCGGCGGCTGCATACGCTTATC 768
DB	721 CTCCTCCGGAAGATGATCCATATGCTGATGATGAATGCGGCGGCTGCATACGCTTATC 780
QY	769 CGGTACTCTGCCATTCGACCAACGAAACATCGCATCGAGCGAGCAGTACTCGGA 828
DB	781 CGGTACTCTGCCATTCGACCAACGAAACATCGCATCGAGCGAGCAGTACTCGGA 840
QY	829 TGGNAGCGGTCTTGTGATCAGATGATCTGGACGAAGAGCATCAGGGGCTCGGCCAG 888
DB	
Db	841 TGGAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGGCCAG 900
QY	889 CCGAATCTTCCCGAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTCTGACCC 948
DB	901 CCGAATCTTCCCGAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTCTGACCC 960
QY	949 ATGGCGATGCTGCTTCCGGAATATCATGTTGGAAATATGGCGGCTTTCTGGATTCATCG 1008
DB	961 ATGGCGATGCTGCTTCCGGAATATCATGTTGGAAATATGGCGGCTTTCTGGATTCATCG 1020
QY	1009 ACTGTGGCCGGTGTGGCGACCGCTATCAGGACATAGCGTTGCTACCCGTGATA 1068
DB	1021 ACTGTGGCCGGTGTGGCGACCGCTATCAGGACATAGCGTTGCTACCCGTGATA 1080
QY	1069 TTGCTGAAGACTTGGGGCGAATGGGCTGACCGCTTCTCTGCTTTACGSTATCGCG 1128
DB	1081 TTGCTGAAGACTTGGGGCGAATGGGCTGACCGCTTCTCTGCTTTACGSTATCGCG 1140
QY	1129 CTCCGATTCGAGCGGATCGCTTCTATCGCTTCTTTCAGAGTCTTCTGAGTTTAAA 1188
DB	1141 CTCCGATTCGAGCGGATCGCTTCTATCGCTTCTTTCAGAGTCTTCTGAGTTTAAA 1200
QY	1189 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248
DB	1201 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1260
QY	1249 AACGTTACTGCCGAAGCGCTTGGAAATAGCGCGGTGTGGTGTGTATATGTTATTT 1308
DB	1261 AACGTTACTGCCGAAGCGCTTGGAAATAGCGCGGTGTGGTGTGTATATGTTATTT 1320
QY	1309 TCCACCATATTCGCGTCTTTTGGCAATGAGGCGCGGAAACCTGCGCCCTCTCTCTTG 1368
DB	1321 TCCACCATATTCGCGTCTTTTGGCAATGAGGCGCGGAAACCTGCGCCCTCTCTCTTG 1380
QY	1369 ACAGGACTTCTAGGGGTCTTTCCCTCTCCGCAAGAAATGCAAGTCTGTGAATGTC 1428
DB	1381 ACAGGACTTCTAGGGGTCTTTCCCTCTCCGCAAGAAATGCAAGTCTGTGAATGTC 1440
QY	1429 GTGAAGAGCAGTTCCTCTGGAAGTCTTTGAAGACAAACCAACGCTGTAGCAACCTT 1488
DB	1441 GTGAAGAGCAGTTCCTCTGGAAGTCTTTGAAGACAAACCAACGCTGTAGCAACCTT 1500
QY	1489 TGCAAGCAGCGGACCCCGCCACCTGGCGACAGTCTCTCGGCGCAAAAGCCAGGTGTA 1548
DB	1501 TGCAAGCAGCGGACCCCGCCACCTGGCGACAGTCTCTCGGCGCAAAAGCCAGGTGTA 1560
QY	1549 TAAGATACCTGCAAGGGCGCAACCCCGAGTCCAGCTTGTGAGTTGATAGTTGTG 1608
DB	1561 TAAGATACCTGCAAGGGCGCAACCCCGAGTCCAGCTTGTGAGTTGATAGTTGTG 1620
QY	1609 GAAAGATCAATGGTCTCTCAAGCGTATTCACAAAGGGGCTGAAGGATGCCAGAG 1668
DB	1621 GAAAGATCAATGGTCTCTCAAGCGTATTCACAAAGGGGCTGAAGGATGCCAGAG 1680
QY	1669 GTACCCCATGTATGGATCTGATCTGGGCGCTCGGTGCAATCTTTACATGTTTGTAG 1728
DB	1681 GTACCCCATGTATGGATCTGATCTGGGCGCTCGGTGCAATCTTTACATGTTTGTAG 1740
QY	1729 TCGAGGTTAAAAACGCTAGGGCCCCCGAACCAAGGGGAGCTGTTTCTTTTGAATAA 1788
DB	1741 TCGAGGTTAAAAACGCTAGGGCCCCCGAACCAAGGGGAGCTGTTTCTTTTGAATAA 1800
QY	1789 CAGATTAATACCATGGCGCTATTCAGCGCTATTCCTCAACAGAGCGAGGCTACTTGGC 1848
DB	1801 CAGATTAATACCATGGCGCTATTCAGCGCTATTCCTCAACAGAGCGAGGCTACTTGGC 1860
QY	1849 TGCATCATCTAGCTCAGCGCGGACAGAAACAGGTTCAGGGGGAGGTCCCAAGTG 1908
DB	1861 TGCATCATCTAGCTCAGCGCGGACAGAAACAGGTTCAGGGGGAGGTCCCAAGTG 1920
QY	1909 GTCTCCACCGCAACAAATCTTTCTTGGCGACCTTGTGCTCAATGCGCTGTGTGACTGTC 1968
DB	1921 GTCTCCACCGCAACAAATCTTTCTTGGCGACCTTGTGCTCAATGCGCTGTGTGACTGTC 1980



QY 1969 TATCATGTGCGGCTCAAAAGACCTTTCGGCCCAAAAGGGCCCCAATCACCCCAATGTAC 2028  
Db 1981 TATCATGTGCGGCTCAAAAGACCTTTCGGCCCAAAAGGGCCCCAATCACCCCAATGTAC 2040  
QY 2029 ACCAATGTGACAGGACCTGTGGCTGGCAAGCGCCCCCGGGCGGTTTCCTTGACA 2088  
Db 2041 ACCAATGTGACAGGACCTGTGGCTGGCAAGCGCCCCCGGGCGGTTTCCTTGACA 2100  
QY 2089 CCATGCACCTGCGGAGCTCGGACCTTTACTTTGGTCAAGGACATGCCGATGCTCATTCG 2148  
Db 2101 CCATGCACCTGCGGAGCTCGGACCTTTACTTTGGTCAAGGACATGCCGATGCTCATTCG 2160  
QY 2149 GTGGCCCGGGGGGCGACAGAGGGGAGCCTACTCTCCCGAGCGCGCTCTCCTACTTG 2208  
Db 2161 GTGGCCCGGGGGGCGACAGGAGGGGAGCCTACTCTCCCGAGCGCGCTCTCCTACTTG 2220  
QY 2209 AAGGCTCTTTCGGGGGCTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCACTTTTCGG 2268  
Db 2221 AAGGCTCTTTCGGGGGCTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCACTTTTCGG 2280  
QY 2269 GGTGCGGTGTGCACCGAGGGGTTTCGGAAGGGGTGGAACTTTGTATACCGTTCGAGTCTATG 2328  
Db 2281 GGTGCGGTGTGCACCGAGGGGTTTCGGAAGGGGTGGAACTTTGTATACCGTTCGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGTCCCGGTCTTACAGGCAAACTGCTGCTCCCTCGGCGGTACCGGAG 2388  
Db 2341 GAAACCACTATGCGGTCCCGGTCTTACAGGCAAACTGCTGCTCCCTCGGCGGTACCGGAG 2400  
QY 2389 ACATTTCCAGGTGGCCCATCTACAGCCCTACTGCTAGCGGCAAGACATAAGGTGCG 2448  
Db 2401 ACATTTCCAGGTGGCCCATCTACAGCCCTACTGCTAGCGGCAAGACATAAGGTGCG 2460  
QY 2449 GGTGCGGTATGACGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCGGTCCCGCCACC 2508  
Db 2461 GGTGCGGTATGACGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCGGTCCCGCCACC 2520  
QY 2509 CTAGGTTTCGGGGGCTATATGCTTAAGGCACATGCTATCGACCTAAACATAGAACCGGG 2568  
Db 2521 CTAGGTTTCGGGGGCTATATGCTTAAGGCACATGCTATCGACCTAAACATAGAACCGGG 2580  
QY 2569 GTAAGGACCATCACACGGGTGCCCCCATCAGTACTCCACTATGCGCAAGTTCCTTGC 2628  
Db 2581 GTAGGACCATCACACGGGTGCCCCCATCAGTACTCCACTATGCGCAAGTTCCTTGC 2640  
QY 2629 GACGTTGTTGCTCTGGGGGCGCTATGACATATAATATGATGAGTGCCACTCAACT 2688  
Db 2641 GACGTTGTTGCTCTGGGGGCGCTATGACATATAATATGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTGGGCACTCGGCACAGTCTCGACCAAGCGGAGCGCTGGAGCG 2748  
Db 2701 GACTCGACCACTATCTGGGCACTCGGCACAGTCTCGACCAAGCGGAGCGCTGGAGCG 2760  
QY 2749 CCACTCTGCTGCTCGCCACCGTACGCTTCGGGATCGGTCAAGTGCACATCCAAAC 2808  
Db 2761 CCACTCTGCTGCTCGCCACCGTACGCTTCGGGATCGGTCAAGTGCACATCCAAAC 2820  
QY 2809 ATCGAGGAGGTGGCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCG 2868  
Db 2821 ATCGAGGAGGTGGCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCG 2880  
QY 2869 ATCGAGACCATCAAGGGGGGAGGCACTCAATTTCTGCAATTCMAAGAAATGTGAT 2928  
Db 2881 ATCGGACCATCAAGGGGGGAGGCACTCAATTTCTGCAATTCMAAGAAATGTGAT 2940  
QY 2929 GAGTCCGCGGAGCTGTCGGCTCGGACCTCAATGCTAGCATATTACCGGGGCTT 2988  
Db 2941 GAGTCCGCGGAGCTATTCGGCTCGGACCTCAATGCTAGCATATTACCGGGGCTT 3000  
QY 2989 GATGTATCCGTCATACCAACTAGCGGAGAGCTCAATTTGCTGTAGCAACGAGCGCTTAATG 3048  
Db 3001 GATGTATCCGTCATACCAACTAGCGGAGAGCTCAATTTGCTGTAGCAACGAGCGCTTAATG 3060

QY 3049 ACGGCTTTTACCGGGGATTTTCACTCAGTGATCGACTGCAATACATGTGTCAACCAGACA 3108  
Db 3061 ACGGCTTTTACCGGGGATTTTCACTCAGTGATCGACTGCAATACATGTGTCAACCAGACA 3120  
QY 3109 GTGCACTTTCAGCTTGGACCCGACCTTACCATTGAGACGACACCGTGCACAAGACGG 3168  
Db 3121 GTGCACTTTCAGCTTGGACCCGACCTTACCATTGAGACGACACCGTGCACAAGACGG 3180  
QY 3169 GTGTCACTGCTGAGCGGCGGAGGAGGAGTGTGTAGGGGAGGATGGGCATTTACAGTTT 3228  
Db 3181 GTGTCACTGCTGAGCGGCGGAGGAGGAGTGTGTAGGGGAGGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTCCAGAGAACCGGCTTCGGGATGTCGATTCCTCGGTTCTGTGGAGTGCTAT 3288  
Db 3241 GTGACTCCAGAGAACCGGCTTCGGGATGTCGATTCCTCGGTTCTGTGGAGTGCTAT 3300  
QY 3289 GACGGGCTGTGCTTGTGTAGAGCTACGGCCCGCGAGACCTCAGTTAGGTTTGGGGCT 3348  
Db 3301 GACGGGCTGTGCTTGTGTAGAGCTACGGCCCGCGAGACCTCAGTTAGGTTTGGGGCT 3360  
QY 3349 TACTTAAACACACAGGCTGCGGCTTCGCGAGGACCATCTGGAGTCTTGGGAGAGGTC 3408  
Db 3361 TACTTAAACACACAGGCTGCGGCTTCGCGAGGACCATCTGGAGTCTTGGGAGAGGTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGACGACGACGAC 3468  
Db 3421 TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGACGACGACGAC 3480  
QY 3469 AACTTCCCTTACTGCTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTCCACT 3528  
Db 3481 AACTTCCCTTACTGCTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTCCACT 3540  
QY 3529 CCATCGTGGGACCAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGGGCA 3588  
Db 3541 CCATCGTGGGACCAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGGGCA 3600  
QY 3589 AGGCGCTGCTGTATAGGCTGGGAGCGGTTCAAAAGAGGTTTACTACACACACCCATA 3648  
Db 3601 AGGCGCTGCTGTATAGGCTGGGAGCGGTTCAAAAGAGGTTTACTACACACACCCATA 3660  
QY 3649 ACCAAATACATCATGGCATGCAATGCTGGCTGACCTGGAGTCTGTCACGAGCACTGGGTG 3708  
Db 3661 ACCAAATACATCATGGCATGCAATGCTGGCTGACCTGGAGTCTGTCACGAGCACTGGGTG 3720  
QY 3709 CTGTGAGCGGAGTCTGTAGAGCTGTGGCGGCTATGCTGACAAAGGAGGCTGCTG 3768  
Db 3721 CTGTGAGCGGAGTCTGTAGAGCTGTGGCGGCTATGCTGACAAAGGAGGCTGCTG 3780  
QY 3769 ATTGTGGCGGAGTCACTTGTCCGAAAGCGGCGCATCATTCGCGACAGGAAAGTCTT 3828  
Db 3781 ATTGTGGCGGAGTCACTTGTCCGAAAGCGGCGCATCATTCGCGACAGGAAAGTCTT 3840  
QY 3829 TACCGGGAGTTTCAATGAGATGGAAGTGGCTGCGCTCACACCTCCCTTACATCGAAACAGGA 3888  
Db 3841 TACCGGGAGTTTCAATGAGATGGAAGTGGCTGCGCTCACACCTCCCTTACATCGAAACAGGA 3900  
QY 3889 ATGCACTGCGGAAACAAATTCAAAAGAGGCAATTCGGGTTGCTGCAACAGCCACCAAG 3948  
Db 3901 ATGCACTGCGGAAACAAATTCAAAAGAGGCAATTCGGGTTGCTGCAACAGCCACCAAG 3960  
QY 3949 GAAAGGAGGCTGCTGCTCCCGTGGGATCCAAAGTGGGAGCCCTCGAAGCTTCTGG 4008  
Db 3961 GAAAGGAGGCTGCTGCTCCCGTGGGATCCAAAGTGGGAGCCCTCGAAGCTTCTGG 4020  
QY 4009 GCGAAGCATATGTGAAATTTTCAATGAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGAAATTTTCAATGAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTGCAACCCCGGAGTACATCTGATGGCAATTCAGAGCTCTATACAGCCCGCTC 4128  
Db 4081 CTGCAACCCCGGAGTACATCTGATGGCAATTCAGAGCTCTATACAGCCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTGTTTAACTCTCTGGGGGATGGTGGCCGCCCACTTGTCT 4188



Db 4141 ACCACCAACATACCTCTGTTAAACATCTGGGGGATGGTGGCCGCCAACTTGCT 4200  
Qy CCTCCAGCGTGTCTTCGCTTTGGTAGGCGCGCATCGCTGGAGCGGCTGTTGGCAGC 4248  
Db 4201 CCTCCAGCGTGTCTTCGCTTTGGTAGGCGCGCATCGCTGGAGCGGCTGTTGGCAGC 4260  
Qy 4249 ATAGGCTTTGGGAAGTCTTGTTGGATATTTTGGCAGGTATGGAGCAGGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTTGGGAAGTCTTGTTGGATATTTTGGCAGGTATGGAGCAGGGGTGGCAGGC 4320  
Qy 4309 GCGTCTGTGGGCTTTAAGGTCAAGAGCGCGAGATGCCCTCAACGAGGACTGGTTAAC 4368  
Db 4321 GCGTCTGTGGGCTTTAAGGTCAAGAGCGCGAGATGCCCTCAACGAGGACTGGTTAAC 4380  
Qy 4369 CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCTCGGGGTGCTGTGGCAGCAGATA 4428  
Db 4381 CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCTCGGGGTGCTGTGGCAGCAGATA 4440  
Qy 4429 CTGCGTCGGCACGTGGGCCACAGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGCGTCGGCACGTGGGCCACAGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCG 4500  
Qy 4489 TTGCGTTCGGGGTAAACAGTCTCCCGCAGCGACATATGTGCGCTGAGAGCGACGCTGCA 4548  
Db 4501 TTGCGTTCGGGGTAAACAGTCTCCCGCAGCGACATATGTGCGCTGAGAGCGACGCTGCA 4560  
Qy 4549 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC 4608  
Db 4561 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC 4620  
Qy 4609 CAGTGGATCAACGAGGACTGTCTCCAGGCCATGCTCGGCTCGTGGCTAAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACGAGGACTGTCTCCAGGCCATGCTCGGCTCGTGGCTAAGAGATGTTGG 4680  
Qy 4669 GATTGGATATGCAAGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCGCGG 4728  
Db 4681 GATTGGATATGCAAGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCGCGG 4740  
Qy 4729 CGATTGCGGGAGTCCCTCTCTCTATGTCACAGTGGGTACAGGGAGTCTGGCGGGG 4788  
Db 4741 CGATTGCGGGAGTCCCTCTCTCTATGTCACAGTGGGTACAGGGAGTCTGGCGGGG 4800  
Qy 4789 GACGGCATATGCAAAACACCTGCCCATTGTGGAGCAGAGTACCGGACATGTGAAAAAC 4848  
Db 4801 GACGGCATATGCAAAACACCTGCCCATTGTGGAGCAGAGTACCGGACATGTGAAAAAC 4860  
Qy 4849 GGTTCATAGAGATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAATTCCTCC 4908  
Db 4861 GGTTCATAGAGATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAATTCCTCC 4920  
Qy 4909 ATTAACGGGTACACACGGGCCCTGACAGCCCTCCCGGCCCAAAATTTCTAGGGCG 4968  
Db 4921 ATTAACGGGTACACACGGGCCCTGACAGCCCTCCCGGCCCAAAATTTCTAGGGCG 4980  
Qy 4969 CTGTGGGGGTGGTGTGAGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGGGGTGGTGTGAGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC 5040  
Qy 5029 GTGACGGCATGACCACTGACAAAGTAAAGTCCCGTGTAGGTTCCGGCCCCCGAATTC 5088  
Db 5041 GTGACGGCATGACCACTGACAAAGTAAAGTCCCGTGTAGGTTCCGGCCCCCGAATTC 5100  
Qy 5089 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGCTCCAGGTCGCAAAACCTCTCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGCTCCAGGTCGCAAAACCTCTCTA 5160  
Qy 5149 CGGAGAGGTGCAATTCCTGGTCCGGCTCAATCAATACCTGTGGGTACAGCTCCCA 5208  
Db 5161 CGGAGAGGTGCAATTCCTGGTCCGGCTCAATCAATACCTGTGGGTACAGCTCCCA 5220  
Qy 5209 TGGAGCCGAACCGGAGTGTGCTACTTCCATGCTCACGACCTCCCACTT 5268

Db 5221 TGGGGGCCGAACCGGACGTAGCAGTGTCTACTTCCATGCTCACCGACCCCTCCACATT 5280  
Qy 5269 ACGGGGAGACGGCTAAGCGTAGCGTGGCAGGGGATCTCCCGCTCTCTGGCCAGCTCA 5328  
Db 5281 ACGGGGAGACGGCTAAGCGTAGCGTGGCAGGGGATCTCCCGCTCTCTGGCCAGCTCA 5340  
Qy 5329 TCAGTACGCCAGCTGTCTGGCCCTTCTTTGAAGGCAACATCACTACCCGTCATGACTCC 5388  
Db 5341 TCAGTACGCCAGCTGTCTGGCCCTTCTTTGAAGGCAACATCACTACCCGTCATGACTCC 5400  
Qy 5389 CCGAGCGTGACCTCATCGAGGCCAACCTCTGTGGGGCAGGAGATGGGGGGAACATC 5448  
Db 5401 CCGAGCGTGACCTCATCGAGGCCAACCTCTGTGGGGCAGGAGATGGGGGGAACATC 5460  
Qy 5449 ACCCGCTGGAGTCAGAAATTAAGTAGTATTTTGGACTCTTTTCGAGCCCTCCAAGCG 5508  
Db 5461 ACCCGCTGGAGTCAGAAATTAAGTAGTATTTTGGACTCTTTTCGAGCCCTCCAAGCG 5520  
Qy 5509 GAGGAGATGAGAGGGAGTATCCGTTCCGGCGGAGATCTCTGGGAGGTCCAGGAAATTC 5568  
Db 5521 GAGGAGATGAGAGGGAGTATCCGTTCCGGCGGAGATCTCTGGGAGGTCCAGGAAATTC 5580  
Qy 5569 CCTCGAGCATGCCCCATATGGGCAAGCCCGGANTPACAACTCTCACTGTTAGAGTCTG 5628  
Db 5581 CCTCGAGCATGCCCCATATGGGCAAGCCCGGANTPACAACTCTCACTGTTAGAGTCTG 5640  
Qy 5629 AAGGACCGGACTACGCTCCCTCAGTGTACACGGGTGTCATTTCCGCTCCCAAGGCC 5688  
Db 5641 AAGGACCGGACTACGCTCCCTCAGTGTACACGGGTGTCATTTCCGCTCCCAAGGCC 5700  
Qy 5689 CCTCCGATACACACTCCAGGAGGAGGACGCTTGTCTGTGTCAGAAATCTACCGTGTCT 5748  
Db 5701 CCTCCGATACACACTCCAGGAGGAGGAGCGGTTGTCTGTGTCAGAAATCTACCGTGTCT 5760  
Qy 5749 TCTGCTTGGCGAGTCTGCCCAAAAGACCTTCGGGAGTCCGAAATCTGTCGGCGTGCAC 5808  
Db 5761 TCTGCTTGGCGAGTCTGCCCAAAAGACCTTCGGGAGTCCGAAATCTGTCGGCGTGCAC 5820  
Qy 5809 AGCGGACGGCAACGGGCTCTCTGACAGCCCTCCGACGCGGAGCGCGGAGTCCGAC 5868  
Db 5821 AGCGGACGGCAACGGGCTCTCTGACAGCCCTCCGACGCGGAGCGCGGAGTCCGAC 5880  
Qy 5869 GTTGAGTCTGACTCTCTCCATGCCCCCTTGAAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTGACTCTCTCCATGCCCCCTTGAAGGGGAGCGGGGATCCCGATCTCAGC 5940  
Qy 5929 GACGGTCTTGTGTCTACCGTAAGCGAGGAGCTAGTGAAGACGTCGTCTGTCTGATG 5988  
Db 5941 GACGGTCTTGTGTCTACCGTAAGCGAGGAGCTAGTGAAGACGTCGTCTGTCTGATG 6000  
Qy 5989 TCCTACACATGGACAGCGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC 6048  
Db 6001 TCCTACACATGGACAGCGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC 6060  
Qy 6049 ATCAATCACTGAGCAACTCTTTGCTCCGTACACCAACTTGGTCTATGCTACAACTCT 6108  
Db 6061 GTCAATCACTGAGCAACTCTTTGCTCCGTACACCAACTTGGTCTATGCTACAACTCT 6120  
Qy 6109 GCGAGCGCAAGCTCTGGGAGAGAGAGTCACTTTGACAGACTGAGGTCCTGGAGCAGC 6168  
Db 6121 GCGAGCGCAAGCTCTGGGAGAGAGAGTCACTTTTACAGACTGAGGTCCTGGAGCAGC 6180  
Qy 6169 CACTACCGGAGCTGTCTCAAGGAGATGAAGGAGCGCTCCAGTTAAGGCTAAACTT 6228  
Db 6181 CACTACCGGAGCTGTCTCAAGGAGATGAAGGAGCGCTCCAGTTAAGGCTAAACTT 6240  
Qy 6229 CTATCCGTGGAGGAGCTGTAAAGTGAACCCCCCACTTCGGCCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGGAGGAGCTGTAAAGTGAACCCCCCACTTCGGCCAGATCTAAATTTGGC 6300  
Qy 6289 TATGGGCAAGAGAGCTGTGGAACTTCCAGAGCGCGCTTAACCAATCCGCTCCGTG 6348  
Db 6301 TATGGGCAAGAGAGCTGTGGAACTTCCAGAGCGCGCTTAACCAATCCGCTCCGTG 6360



Db 1141 CTCCGATTTCGACGGGATCGCCTTCTATCGCCTTCTTGAAGAGTTCTTCTGAGTTTAAA 1200  
QY 1189 CAGACCAACACGGTTTCCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCGCCCCCT 1248  
Db 1201 CAGACCAACACGGTTTCCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCGCCCCCT 1260  
QY 1249 AACGTTACTGCGGACGAGCGCTTGGATAAGCGCGGTGTGGTTGTCTATATGTTATTT 1308  
Db 1261 AACGTTACTGCGGACGAGCGCTTGGATAAGCGCGGTGTGGTTGTCTATATGTTATTT 1320  
QY 1309 TCCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAACCTGGGCCCTGTCTCTTG 1368  
Db 1321 TCCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAACCTGGGCCCTGTCTCTTG 1380  
QY 1369 ACAGAGATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGAATGTC 1428  
Db 1381 ACAGAGATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGAATGTC 1440  
QY 1429 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAACCAAGCTCTGTAGCACCTTT 1488  
Db 1441 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAACCAAGCTCTGTAGCACCTTT 1500  
QY 1489 TGCAGGACGAGGAAACCCCTCGGACAGGTGCTCTCGGCCAAAGCCACGTGTA 1548  
Db 1501 TGCAGGACGAGGAAACCCCTCGGACAGGTGCTCTCGGCCAAAGCCACGTGTA 1560  
QY 1549 TAAGATACCTGCAAGAGGCGCAACCCAGTGCACGTTGTAGTTGATGTTGTG 1608  
Db 1561 TAAGATACCTGCAAGAGGCGCAACCCAGTGCACGTTGTAGTTGATGTTGTG 1620  
QY 1609 GAAAGATCAAAATGGCTCTCTCAAGCGTATTCACAGGGGCTGAAGATGCCAGAG 1668  
Db 1621 GAAAGATCAAAATGGCTCTCTCAAGCGTATTCACAGGGGCTGAAGATGCCAGAG 1680  
QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCAATGTTTACATGTTTGTAG 1728  
Db 1681 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCAATGTTTACATGTTTGTAG 1740  
QY 1729 TCGAGGTTAAACACGTTCTAGGCCCCCGGAAACCAAGGGGAGCTGGTTTTCTTTGAAAA 1788  
Db 1741 TCGAGGTTAAACACGTTCTAGGCCCCCGGAAACCAAGGGGAGCTGGTTTTCTTTGAAAA 1800  
QY 1789 CAGGATTAATACCTAGCGGCTTATTACGCTTACTCCCAACAGAGCGGAGGCTACTTGGC 1848  
Db 1801 CAGGATTAATACCTAGCGGCTTATTACGCTTACTCCCAACAGAGCGGAGGCTACTTGGC 1860  
QY 1849 TGCATCATCTACTAGCTTCACAGGCGGGACAGGAACCAAGTTCGAGGGGAGGTCACAGTG 1908  
Db 1861 TGCATCATCTACTAGCTTCACAGGCGGGACAGGAACCAAGTTCGAGGGGAGGTCACAGTG 1920  
QY 1909 GTCTCCACCGGAACACAAATCTTTCTGGGACCTGCGTCAATGCGGTGTGTGGACTGTGTC 1968  
Db 1921 GTCTCCACCGGAACACAAATCTTTCTGGGACCTGCGTCAATGCGGTGTGTGGACTGTGTC 1980  
QY 1969 TATCATGTGCGGCTCAAAGACCTTTCCGCGCCCAAGGGCCCAATCACCCAAATGTAC 2028  
Db 1981 TATCATGTGCGGCTCAAAGACCTTTCCGCGCCCAAGGGCCCAATCACCCAAATGTAC 2040  
QY 2029 ACCAATGTGAACAGGACCTCGTGGTGGCAAGCGCCCCCGGGGCGGCTTCCTTGACA 2088  
Db 2041 ACCAATGTGAACAGGACCTCGTGGTGGCAAGCGCCCCCGGGGCGGCTTCCTTGACA 2100  
QY 2089 CCATGACCTGCGGACGCTCGGACCTTTACTTGGTCAAGGAGGATGCGGATGTATCCG 2148  
Db 2101 CCATGACCTGCGGACGCTCGGACCTTTACTTGGTCAAGGAGGATGCGGATGTATCCG 2160  
QY 2149 GTGCGCGGCGGCGGACAGCAGGGGAGCTACTCTCCCGCAGGCGGCTCTCTACTTGT 2208  
Db 2161 GTGCGCGGCGGCGGACAGCAGGGGAGCTACTCTCTCCCGCAGGCGGCTCTCTACTTGT 2220  
QY 2209 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCCCTCGGGGACAGCTGTGGGCACTTTTCG 2268  
Db 2221 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCCCTCGGGGACAGCTGTGGGCACTTTTCG 2280

Db 61 TCTTACCGAGAAAGCGTTAGCCATAGCGTTAGTATAGTGTGTCGAGCCTCAGGAC 120  
QY 121 CCCCTCTCCGGGAGAGCATAAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGAG 180  
Db 121 CCCCTCTCCGGGAGAGCATAAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGAG 180  
QY 181 GACGACCGGTCCTTTCTTGTGATCAACCGGCTCAATGCTCGGAGATTTGGGGTGCCTCC 240  
Db 181 GACGACCGGTCCTTTCTTGTGATCAACCGGCTCAATGCTCGGAGATTTGGGGTGCCTCC 240  
QY 241 GCGAGACTGTAGCGGAGTAGTGTGGTTCGGAAGGCTTTGGTACTGCTGATAGG 300  
Db 241 GCGAGACTGTAGCGGAGTAGTGTGGTTCGGAAGGCTTTGGTACTGCTGATAGG 300  
QY 301 GTGCTTGGAGTGCCCGGAGGCTCTGTAGACCGTGCACATGAGCAGAAATCCTAAAC 360  
Db 301 GTGCTTGGAGTGCCCGGAGGCTCTGTAGACCGTGCACATGAGCAGAAATCCTAAAC 360  
QY 361 CTAAAGAAAAACAAA-----GGGCGGCCATGATTGAACAAGATGGATTGC 408  
Db 361 CTAAAGAAAAACAAAAGCTAACCAACGGCGGCCCATGATTGAACAAGATGGATTGC 420  
QY 409 ACAGAGTTCTCCGCGCTTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAAGA 468  
Db 421 ACAGAGTTCTCCGCGCTTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAAGA 480  
QY 469 CAATCGGCTGTCTGATGCGCGCTGTTCGCGTGTGACGCGAGGGCGCGGTTCTTT 528  
Db 481 CAATCGGCTGTCTGATGCGCGCTGTTCGCGTGTGACGCGAGGGCGCGGTTCTTT 540  
QY 529 TTGTCAAGACGACCTGTCGCGTCCCTGAATGAATGCAAGTGCAGGACGAGCAGCGGCTAT 588  
Db 541 TTGTCAAGACGACCTGTCGCGTCCCTGAATGAATGCAAGTGCAGGACGAGCAGCGGCTAT 600  
QY 589 CGTGGCTGGCCACGACGGGCTTCTTTCGCGAGTGTGCTCGAAGTGTCTACTGAAGCGG 648  
Db 601 CGTGGCTGGCCACGACGGGCTTCTTTCGCGAGTGTGCTCGAAGTGTCTACTGAAGCGG 660  
QY 649 GAAGGACTGTGCTGTTTGGGGAAGTGCAGGCGGAGGATCTCTGTCATCTCACTTGTG 708  
Db 661 GAAGGACTGTGCTGTTTGGGGAAGTGCAGGCGGAGGATCTCTGTCATCTCACTTGTG 720  
QY 709 CTCTGCGGAGAAAGTATCCATATGCTGATGAATGCGCGGCTGCATACGCTTGTATC 768  
Db 721 CTCTGCGGAGAAAGTATCCATATGCTGATGAATGCGCGGCTGCATACGCTTGTATC 780  
QY 769 CGGCTACTGCGCATTCGACCAACCAAGCGAAATCGCATCGAGCGAGCAGTACTCGGA 828  
Db 781 CGGCTACTGCGCATTCGACCAACCAAGCGAAATCGCATCGAGCGAGCAGTACTCGGA 840  
QY 829 TGGAGCGGCTGTGTCGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAG 888  
Db 841 TGGAGCGGCTGTGTCGATCAGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAG 900  
QY 889 CCGAATCTGTCGCGAGGCTCAAGCGGCGCATGCCCCGAGCGAGATCTGTCGTGACCC 948  
Db 901 CCGAATCTGTCGCGAGGCTCAAGCGGCGCATGCCCCGAGCGAGATCTGTCGTGACCC 960  
QY 949 ATGCGGATGCTGTTGCGGAATATGTTGGGAAATGCGCGCTTTTCTGGATTATCG 1008  
Db 961 ATGCGGATGCTGTTGCGGAATATGTTGGGAAATGCGCGCTTTTCTGGATTATCG 1020  
QY 1009 ACTGTGCGGCTGCGGTGTCGCGGCTATCAGGACATAGCGTTGCGTACCGGTGATA 1068  
Db 1021 ACTGTGCGGCTGCGGTGTCGCGGACCGCTATCAGGACATAGCGTTGCGTACCGGTGATA 1080  
QY 1069 TTGTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTTTTACGGTATCGCG 1128  
Db 1081 TTGTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTTTTACGGTATCGCG 1140  
QY 1129 CTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTGAGAGTCTTCTGAGTTTAAA 1188

QY	2269	GCTGCCGTGTGCACCCGAGGGGTTCGAAAGCCGGTGGACTTTGTACCCGTCGAGTCTATG	2322
Db	2281	GCTGCCGTGTGCACCCGAGGGGTTCGAAAGCCGGTGGACTTTGTACCCGTCGAGTCTATG	2340
QY	2329	GAAAACCACTATGCGGTCCCGGTCCTTACGGCAAACTCGTCCCTCGGCGCGTACCGCAG	2388
Db	2341	GAAAACCACTATGCGGTCCCGGTCCTTACGGCAAACTCGTCCCTCGGCGCGTACCGCAG	2400
QY	2389	ACATTCAGGTGGCCCATCTACACGCGCCCTACCTGGTAGCGGCAAGAGCACTAAGGTGCCG	2448
Db	2401	ACATTCAGGTGGCCCATCTACACGCGCCCTACCTGGTAGCGGCAAGAGCACTAAGGTGCCG	2460
QY	2449	GCTGCGTATCGACCCAAAGGATATAAGTGCTTGTCTGTGAACCGTCCGTCGCGCCAC	2508
Db	2461	GCTGCGTATCGACCCAAAGGATATAAGTGCTTGTCTGTGAACCGTCCGTCGCGCCAC	2520
QY	2509	CTAGGTTTCGGGGGTATATGTCTAAGGCACATGTTATCGACCTTAACATCATCAAAACCGG	2568
Db	2521	CTAGGTTTCGGGGGTATATGTCTAAGGCACATGTTATCGACCTTAACATCATCAAAACCGG	2580
QY	2569	GTAAGGACCATCACACGGGTGCCCATCACGTACTCCACCTATGCGAAAGTTCTTGCC	2628
Db	2581	GTAGGNACATCACACGGGTGCCCATCACGTACTCCACCTATGCGAAAGTTCTTGCC	2640
QY	2629	GACGGTGGTCTCTGGGGCGCCTATGACATCATATAATGTGATGAGTGCCACTCAACT	2688
Db	2641	GACGGTGGTCTCTGGGGCGCCTATGACATCATATAATGTGATGAGTGCCACTCAACT	2700
QY	2689	GACTCGACACTATCTGGGCATCGGCACAGTCTGTGACCAAGCGGAGACGGCTGAGCG	2748
Db	2701	GACTCGACACTATCTGGGCATCGGCACAGTCTGTGACCAAGCGGAGACGGCTGAGCG	2760
QY	2749	CGACTCGTGTCTGCGCACCGCTACGCCCTCCGGGATCGGTACCGTGCACATCCAAAC	2808
Db	2761	CGACTCGTGTCTGCGCACCGCTACGCCCTCCGGGATCGGTACCGTGCACATCCAAAC	2820
QY	2809	ATCGAGAGTGGCTGTCTCCAGACTTGGAGAAATCCCTTTATGGCAAAGCCATCCCC	2868
Db	2821	ATCGAGAGTGGCTGTCTCCAGACTTGGAGAAATCCCTTTATGGCAAAGCCATCCCC	2880
QY	2869	ATCGAGACATCAAGGGGGGAGCAGCTCATTTTCTGGCATTCGAAGAAGAAATGTGAT	2928
Db	2881	ATCGAGACATCAAGGGGGGAGCAGCTCATTTTCTGGCATTCGAAGAAGAAATGTGAT	2940
QY	2929	GAGCTCGCCGAGCTGCCGCTCGAGCTCAATGCTGTAGCATATACCGGGCCCT	2988
Db	2941	GAGCTCGCCGAGCTATCCGGCTCGAGCTCAATGCTGTAGCATATACCGGGCCCT	3000
QY	2989	GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTCTGTAGCAACGGACGCTTAATG	3048
Db	3001	GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTCTGTAGCAACGGACGCTTAATG	3060
QY	3049	ACGGCTTTACCGCGATTCGACTCAGTGATCGACTGGAATACATGTGTCAACCGACACA	3108
Db	3061	ACGGCTTTACCGCGATTCGACTCAGTGATCGACTGGAATACATGTGTCAACCGACACA	3120
QY	3109	GTGACTTACGCTGGACCGACCTTACCATTCAGACGACCGTGCACACGAGCGG	3168
Db	3121	GTGACTTACGCTGGACCGACCTTACCATTCAGACGACCGTGCACACGAGCGG	3180
QY	3169	GTGTACGCTCGACCGCGGAGGAGGACTGGTAGGGGAGGATGGGCATTTACAGGTTT	3228
Db	3181	GTGTACGCTCGACCGCGGAGGAGGACTGGTAGGGGAGGATGGGCATTTACAGGTTT	3240
QY	3229	GTGACTCCAGAGAAACCGCCTCTGGGCATGTTCGATTCCTCGGTCGTGCGAGTGCTAT	3288
Db	3241	GTGACTCCAGAGAAACCGCCTCTGGGCATGTTCGATTCCTCGGTCGTGCGAGTGCTAT	3300
QY	3289	GACGCGGCTGTGCTTGGTAGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCCGGCT	3348
Db	3301	GACGCGGCTGTGCTTGGTAGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCCGGCT	3360

Qy	3349	TACCTAAACACACACAGGGTTCGCCGCTCTGCCAGGACCAATCTCGAGTTCTGGGAGAGCGTC	3408
Db	3361	TACCTAAACACACACAGGGCTGCCCGTCTGCCAGGACCAATCTCGAGTTCTGGGAGAGCGTC	3420
Qy	3409	TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCTCCAGACTAAGCAGGCAGAGAGAC	3468
Db	3421	TTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCTCCAGACTAAGCAGGCAGAGAGAC	3480
Qy	3469	AAC TTCCTCCCTACCTGGTAGCATACCAAGGCTACGGTGTGGCCAGGGCTCAGGCTCCACCT	3528
Db	3481	AAC TTCCTCCCTACCTGGTAGCATACCAAGGCTACGGTGTGGCCAGGGCTCAGGCTCCACCT	3540
Qy	3529	CCATCTGTGGACCAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCCA	3588
Db	3541	CCATCTGTGGACCAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCCA	3600
Qy	3589	ACGCCCTGCTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCAACACACACCCCATTA	3648
Db	3601	ACGCCCTGCTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCAACACACCCCATTA	3660
Qy	3649	ACCAATATACATCATGGCATCATGTCCGCTGACCTCGAGGTCGTCTACGAGCACTGGGTG	3708
Db	3661	ACCAATATACATCATGGCATCATGTCCGCTGACCTCGAGGTCGTCTACGAGCACTGGGTG	3720
Qy	3709	CTGGTAGGCGGAGTCTTAGCAGCTCTGGCCGCGTATTGTCCTTGACAAACAGGACGCTGGTC	3768
Db	3721	CTGGTAGGCGGAGTCTTAGCAGCTCTGGCCGCGTATTGTCCTTGACAAACAGGACGCTGGTC	3780
Qy	3769	ATTGTGGGAGGATCATCTTGTCGGAAAGCCGGCCATCATTCCTCCGACAGGGAAGTCCTT	3828
Db	3781	ATTGTGGGAGGATCATCTTGTCGGAAAGCCGGCCATCATTCCTCCGACAGGGAAGTCCTT	3840
Qy	3829	TACCGGAGTTCGATGAGATGGAAGTAGTGGCCTCACCTCCCTTACATTCGAACACAGGA	3888
Db	3841	TACCGGAGTTCGATGAGATGGAAGTAGTGGCCTCACCTCCCTTACATTCGAACACAGGA	3900
Qy	3889	ATGCAGCTCCCGAAACAAATTCGGAAGGCAATCGGGTGTCTGMAACAGGCACCAAG	3948
Db	3901	ATGCAGCTCCCGAAACAAATTCGGAAGGCAATCGGGTGTCTGMAACAGGCACCAAG	3960
Qy	3949	CAAGCAGAGCTGCTGCTCCCGTGGTGGATCCAAAGTGGCGGACCTCGAAGCCTTCTGG	4008
Db	3961	CAAGCAGAGCTGCTGCTCCCGTGGTGGATCCAAAGTGGCGGACCTCGAAGCCTTCTGG	4020
Qy	4009	GCGAAGCATATGTGGAATTTTCATCAGCGGATACAAATATTAGCAGGCTGTCCACTCTG	4068
Db	4021	GCGAAGCATATGTGGAATTTTCATCAGCGGATACAAATATTAGCAGGCTGTCCACTCTG	4080
Qy	4069	CCTGGCAACCCCGATAGCATCACTGATGGCATTCACAGCCTTATACCAAGCCCGCTC	4128
Db	4081	CCTGGCAACCCCGATAGCATCACTGATGGCATTCACAGCCTTATACCAAGCCCGCTC	4140
Qy	4129	ACCACCAACATACCCCTCTGTTTAACTCCTGGGGGATGGGTGGCCGCGCACTTGCT	4188
Db	4141	ACCACCAACATACCCCTCTGTTTAACTCCTGGGGGATGGGTGGCCGCGCACTTGCT	4200
Qy	4189	CCTCCAGCGCTGCTCTGCTTTCTGTAAGCGCCGGCATCGCTGAGCGGCTGTGGCAGC	4248
Db	4201	CCTCCAGCGCTGCTCTGCTTTCTGTAAGCGCCGGCATCGCTGAGCGGCTGTGGCAGC	4260
Qy	4249	ATAGGCCCTTGGGAAGTGTCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGC	4308
Db	4261	ATAGGCCCTTGGGAAGTGTCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGC	4320
Qy	4309	CGGCTCTGGCCTTTAAGGTCATAGCGGGAGATGCCCTCCACGAGGACCTGGTTAAC	4368
Db	4321	CGGCTCTGGCCTTTAAGGTCATAGCGGGAGATGCCCTCCACGAGGACCTGGTTAAC	4380
Qy	4369	CTACTCCTCGTATCCPCTCCCTGGGCCCTAGTCTGTGGGGTCTGTGTCGACAGCAT	4428
Db	4381	CTACTCCTCGTATCCPCTCCCTGGGCCCTAGTCTGTGGGGTCTGTGTCGACAGCAT	4440
Qy	4429	CTGGTCGGCACCTGGGCCCCACGGGAGGGGGCTGTGTCAGTGTGATGAACCGGCTCATAGCG	4488



```
QY 6649 TGTGTTGACTCAACGGTCACTGAGAAATGACATCCCTGTTGAGGAGTCAATCTACCAATGT 6708
Db 6661 TGTGTTGACTCAACGGTCACTGAGAAATGACATCCCTGTTGAGGAGTCAATCTACCAATGT 6720
QY 6709 TGTGACTTGGCCCGGAGACGACAGAGCCATAAGTCCGCTCACAGAGCGGCTTTACATC 6768
Db 6721 TGTGACTTGGCCCGGAGACGACAGAGCCATAAGTCCGCTCACAGAGCGGCTTTACATC 6780
QY 6769 GGGGGCCCCCTGACTAAATTTAAAGGAGAACTCGGCTATCGCGGTTACGCGCGAGC 6828
Db 6781 GGGGGCCCCCTGACTAAATTTAAAGGAGAACTCGGCTATCGCGGTTACGCGCGAGC 6840
QY 6829 GGTGACTGACGACGACGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTCGGCGC 6888
Db 6841 GGTGACTGACGACGACGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTCGGCGC 6900
QY 6889 TGTGAGCTGCGAAGCTCCAGGACTGCAGATGCTCGTATGGGAGACGACCTTCGTT 6948
Db 6901 TGTGAGCTGCGAAGCTCCAGGACTGCAGATGCTCGTATGGGAGACGACCTTCGTT 6960
QY 6949 ATCTGTAAGCGCGGGACCCCAAGAGGACGAGGAGCCTACGGGCTTTACCGAGGCT 7008
Db 6961 ATCTGTAAGCGCGGGACCCCAAGAGGACGAGGAGCCTACGGGCTTTACCGAGGCT 7020
QY 7009 ATGACTAGATCTGTGCCCCCTCGGGGACCGCCCAACCAAGAAATACGACTTGGAGTTG 7068
Db 7021 ATGACTAGATCTGTGCCCCCTCGGGGACCGCCCAACCAAGAAATACGACTTGGAGTTG 7080
QY 7069 ATAAATCATGCTCCTCAATGTGTGAGTGCAGGACGATGCTGCGCAAAAGGTTGATC 7128
Db 7081 ATAAATCATGCTCCTCAATGTGTGAGTGCAGGACGATGCTGCGCAAAAGGTTGATC 7140
QY 7129 TATCTCACCGGTGACCCACCAACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7188
Db 7141 TATCTCACCGGTGACCCACCAACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7200
QY 7189 ACTCAGTCAATCTCTGGCTAGGAAACATCATCATGATGATGCGCCCACTTGTGGGCAAGG 7248
Db 7201 ACTCAGTCAATCTCTGGCTAGGAAACATCATCATGATGATGCGCCCACTTGTGGGCAAGG 7260
QY 7249 ATGATCTGTGATGACTCATTTCTTCCATCTTCCATCTAGCTCAGGAACAACCTTCAAAAGCC 7308
Db 7261 ATGATCTGTGATGACTCATTTCTTCCATCTTCCATCTAGCTCAGGAACAACCTTCAAAAGCC 7320
QY 7309 CTAGATTGTGATGACTCAGGGGCTGTTACTCCATTGAGCCACTTGACCTACCTCAGATC 7368
Db 7321 CTAGATTGTGATGACTCAGGGGCTGTTACTCCATTGAGCCACTTGACCTACCTCAGATC 7380
QY 7369 ATTCAAGACTCCATGGGCTTAGCGGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7428
Db 7381 ATTCAAGACTCCATGGGCTTAGCGGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7440
QY 7429 AATAGGTGGCTTCATGCTCAGGAAACTTGGGTTACCGCCCTTGGAGTCTGGAGACAT 7488
Db 7441 AATAGGTGGCTTCATGCTCAGGAAACTTGGGTTACCGCCCTTGGAGTCTGGAGACAT 7500
QY 7489 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTGTGGC 7548
Db 7501 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTGTGGC 7560
QY 7549 AAGTACTCTTCAACTGGGAGTAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7608
Db 7561 AAGTACTCTTCAACTGGGAGTAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7620
QY 7609 TCCAGTTCGATTAATCCAGCTGTTGTTGTTGTTACAGGGGGAGACATATATAC 7668
Db 7621 TCCAGTTCGATTAATCCAGCTGTTGTTGTTGTTACAGGGGGAGACATATATAC 7680
QY 7669 AGCTGTCTCGTGCCGACCCCGCTGTTTCATGTGTTGCTTACTCTTCTGTTAGGG 7728
Db 7681 AGCTGTCTCGTGCCGACCCCGCTGTTTCATGTGTTGCTTACTCTTCTGTTAGGG 7740
```

```
QY 7729 CTAGCATCTATCTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7788
Db 7741 CTAGCATCTATCTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7800
QY 7789 CCATCCCTCTTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7848
Db 7801 CCATCCCTCTTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7849 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7908
Db 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7909 CCTAGTCAAGCTAGCTGTAAGGTCGTCGAGCGCTTGACTGCAGAGAGTGTGATAC 7968
Db 7921 CCTAGTCAAGCTAGCTGTAAGGTCGTCGAGCGCTTGACTGCAGAGAGTGTGATAC 7980
QY 7969 TGGCCTCTCTGCAGATCAAGT 7989
Db 7981 TGGCCTCTCTGCAGATCAAGT 8001

RESULT 12
LOCUS AR406049 8001 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 28 from patent US 6630343.
ACCESSION AR406049
VERSION AR406049.1 GI:40155176
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 8001)
AUTHORS Bartenschlager,R.
TITLE Hepatitis C virus culture system
JOURNAL Patent: US 6630343-A 28 07-OCT-2003;
FEATURES
source
1. 8001
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.4%; Score 7939.8; DB 6; Length 8001;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 7972; Conservative 0; Mismatches 17; Indels 12; Gaps 1;

QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Db 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
QY 61 TCTTCACGAGAAAGCGTCTAGCCATGGCTTAGTATGAGTGTCTGAGGCTCCAGGAC 120
Db 61 TCTTCACGAGAAAGCGTCTAGCCATGGCTTAGTATGAGTGTCTGAGGCTCCAGGAC 120
QY 121 CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGTGAGTACACCGGAATTCAGGAC 180
Db 121 CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGTGAGTACACCGGAATTCAGGAC 180
QY 181 GACGACCCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTGGAGATTTGGGCGTGCCTCC 240
Db 181 GACGACCCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTGGAGATTTGGGCGTGCCTCC 240
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTTGTGTTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTTGTGTTACTGCTGATAGG 300
QY 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAGACCGTGACCATGAGCAGCATCTTAAC 360
Db 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAGACCGTGACCATGAGCAGCATCTTAAC 360
QY 361 CTCAGAGAAACCAACCAACGTAACCAACCGGCGGCGCCATGATTGAAACAAGATGATTGC 408
Db 361 CTCAGAGAAACCAACCAACGTAACCAACCGGCGGCGCCATGATTGAAACAAGATGATTGC 420
```



409 ACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGTATCACTGGGCACAACAGA 468  
Db |||||  
421 ACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGTATGACTGGGCAACAACAGA 480  
Qy |||||  
469 CAATCGGCTGCTCTGATGCCCGGTTCCGGCTGTACGCCGACAGGGGCGCCGGTCTTTT 528  
Db |||||  
481 CAATCGGCTGCTCTGATGCCCGGTTTCCGGCTGTACGCCGACAGGGGCGCCGGTCTTTT 540  
Qy |||||  
529 TTGTCAGACAGCACTGTCCGGTCCCTGGAATGAATCACTGACAGCAGGCGAGCGGGCTAT 588  
Db |||||  
541 TTGTCAGACAGCACTGTCCGGTCCCTGGAATGAATCACTGACAGCAGGCGAGCGGGCTAT 600  
Qy |||||  
589 CGTGGCTGGCCACAGCAGGGCGTTCTTGGCAGCTGTCTCGAGTGTCTGACTGAAGCGG 648  
Db |||||  
601 CGTGGCTGGCCACAGCAGGGCGTTCTTGGCAGCTGTCTCGAGTGTCTGACTGAAGCGG 660  
Qy |||||  
649 GAAAGGACTGGCTGTATTTGGGCGAAGTTCGGGGCAGGATCTCTGTCATCTCACCTTG 708  
Db |||||  
661 GAAAGGACTGGCTGTATTTGGGCGAAGTTCGGGGCAGGATCTCTGTCATCTCACCTTG 720  
Qy |||||  
709 CTCCTGCCGAAAGTATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTGATC 768  
Db |||||  
721 CTCCTGCCGAAAGTATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTGATC 780  
Qy |||||  
769 CGGCTACTCGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGACGTACTCGGA 828  
Db |||||  
781 CGGCTACTCGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGACGTACTCGGA 840  
Qy |||||  
829 TGAAGCGGCTTGTGATCAGGATGATCTGGAAGCAAGAGCATCAGGGGCTCGGCCAG 888  
Db |||||  
841 TGAAGCGGCTTGTGATCAGGATGATCTGGAAGCAAGAGCATCAGGGGCTCGGCCAG 900  
Qy |||||  
889 CCGAAGTGTTCGCGAGCTCAAGCGCGCATGCGCGGAGGATCTGTCGTGACCC 948  
Db |||||  
901 CCGAAGTGTTCGCGAGCTCAAGCGCGCATGCGCGGAGGATCTGTCGTGACCC 960  
Qy |||||  
949 ATGCGGATGCTGCTGTCCGAATATCATGTGGAAATGCGCGCTTTCTGGAATCATCG 1008  
Db |||||  
961 ATGCGGATGCTGCTGTCCGAATATCATGTGGAAATGCGCGCTTTCTGGAATCATCG 1020  
Qy |||||  
1009 ACTGTGCGCGCTGGGTGTGGCGAACCGCTATACGACATAGCGTGGCTACCGGTGATA 1068  
Db |||||  
1021 ACTGTGCGCGCTGGGTGTGGCGAACCGCTATACGACATAGCGTGGCTACCGGTGATA 1080  
Qy |||||  
1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTAGGGTATCGCG 1128  
Db |||||  
1081 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTAGCGTATCGCG 1140  
Qy |||||  
1129 CTCGGAATTCGACGCGATCGCTTCTATCGCTTCTTTGACGAGTCTTCTGAGTTTAAA 1188  
Db |||||  
1141 CTCGGAATTCGACGCGATCGCTTCTATCGCTTCTTTGACGAGTCTTCTGAGTTTAAA 1200  
Qy |||||  
1189 CAGACACAAGGTTTCCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCGCT 1248  
Db |||||  
1201 CAGACACAAGGTTTCCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCGCT 1260  
Qy |||||  
1249 AACCTTACTGGCGAAGCCCTTGGATTAAGCGCGGTGCGTTGTGCTATATGTTATTT 1308  
Db |||||  
1261 AACCTTACTGGCGAAGCCCTTGGATTAAGCGCGGTGCGTTGTGCTATATGTTATTT 1320  
Qy |||||  
1309 TCACACATATGCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTG 1368  
Db |||||  
1321 TCACACATATGCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTG 1380  
Qy |||||  
1369 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGGAATGTC 1428  
Db |||||  
1381 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGGAATGTC 1440  
Qy |||||  
1429 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTTGAAGACAAACAACGCTGTGAGCACCTT 1488  
Db |||||  
1441 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTTGAAGACAAACAACGCTGTGAGCACCTT 1500  
Qy |||||  
1489 TGCAGGAGCGGAACCCCCACCTGTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGTTGA 1548

1501 TGCAGGAGCGGAACCCCCACCTTGGCGACAGGTGCCTCTCGCGCCAAAAGCCACGTGA 1560  
Qy |||||  
1549 TAAGATACACTGCTCAAAAGGGGCAACAACCCAGTGCACCTGTGTGAGTTGATAGTTGTG 1608  
Db |||||  
1561 TAAGATACACTGCTCAAAAGGGGCAACAACCCAGTGCACCTGTGTGAGTTGATAGTTGTG 1620  
Qy |||||  
1609 GAAAGAGTCAATAGGCTCTCTCAAGCGTATTCAAAAGGGGCTGAAGGATGCCAGAAG 1668  
Db |||||  
1621 GAAAGAGTCAATAGGCTCTCTCAAGCGTATTCAAAAGGGGCTGAAGGATGCCAGAAG 1680  
Qy |||||  
1669 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCATCTTTACATGTGTTTAG 1728  
Db |||||  
1681 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCATCTTTACATGTGTTTAG 1740  
Qy |||||  
1729 TCGAGGTTAAAAAACGCTTAGGCCCCCGCAACACCGGGGAGCTGTTTCTTTGAAAAA 1788  
Db |||||  
1741 TCGAGGTTAAAAAACGCTTAGGCCCCCGCAACACCGGGGAGCTGTTTCTTTGAAAAA 1800  
Qy |||||  
1789 CACGATATACCTAGGCGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1848  
Db |||||  
1801 CACGATATACCTAGGCGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1860  
Qy |||||  
1849 TGCATCATCTAGTACCTCAAGGCGGGACAGGAACACAGGTTCGAGGGGAGGTCCAAGTG 1908  
Db |||||  
1861 TGCATCATCTAGTACCTCAAGGCGGGACAGGAACACAGGTTCGAGGGGAGGTCCAAGTG 1920  
Qy |||||  
1909 GTCTCCACCGCAACACAATCTTTCTCGCGACCTTGCCTCAATGCGGTGTGTGACTGTC 1968  
Db |||||  
1921 GTCTCCACCGCAACACAATCTTTCTCGCGACCTTGCCTCAATGCGGTGTGTGACTGTC 1980  
Qy |||||  
1969 TATCATGTCGCGCTCAAGAGCCCTTTCGCGGCCCAAGGGCCCAATCACCAATGTAC 2028  
Db |||||  
1981 TATCATGTCGCGCTCAAGAGCCCTTTCGCGGCCCAAGGGCCCAATCACCAATGTAC 2040  
Qy |||||  
2029 ACCAATGTGACACGAGACCTCTGCTGCTGCAAGCGCCCGCGGGCGCTTCCTTGACA 2088  
Db |||||  
2041 ACCAATGTGACACGAGACCTCTGCTGCTGCAAGCGCCCGCGGGGCGCTTCCTTGACA 2100  
Qy |||||  
2089 CCATGCACTTGGGCGAGCTCGGACCTTTTACTTGTGTCACGAGGATGCGGATGTCATTCG 2148  
Db |||||  
2101 CCATGCACTTGGGCGAGCTCGGACCTTTTACTTGTGTCACGAGGATGCGGATGTCATTCG 2160  
Qy |||||  
2149 GTGCGCGCGGCGGACAGCAGGGGAGCTACTCTCCCGAGGCGCGTCTCCACTTGT 2208  
Db |||||  
2161 GTGCGCGCGGCGGACAGCAGGGGAGCTACTCTCCCGAGGCGCGTCTCCACTTGT 2220  
Qy |||||  
2209 AAGGGCTTCTTGGGCGGTCCACTGCTCTGCTCGCGGCAACGCTGTGGGCTCTTTTCGG 2268  
Db |||||  
2221 AAGGGCTTCTTGGGCGGTCCACTGCTCTGCTCGCGGCAACGCTGTGGGCTCTTTTCGG 2280  
Qy |||||  
2269 GCTGCGGTGTGCAACCCGAGGGGTTGGAAGGGGTGACCTTTGACCCGTCGAGTCTATG 2328  
Db |||||  
2281 GCTGCGGTGTGCAACCCGAGGGGTTGGAAGGGGTGACCTTTGACCCGTCGAGTCTATG 2340  
Qy |||||  
2329 GAAACCACTATCGGTCCTCCGCTCTTTCACGGAACACTCGTCCCTCCGCGGCTACCGAG 2388  
Db |||||  
2341 GAAACCACTATCGGTCCTCCGCTCTTTCACGGAACACTCGTCCCTCCGCGGCTACCGAG 2400  
Qy |||||  
2389 ACATTCAGGTGGCCCACTCTACACGCGCTTACTGCTAGCGGCAAGAGCACTAAGGTGCGG 2448  
Db |||||  
2401 ACATTCAGGTGGCCCACTCTACACGCGCTTACTGCTAGCGGCAAGAGCACTAAGGTGCGG 2460  
Qy |||||  
2449 GCTGCGGTATGCAACCCGAGGGTATAGGTGCTTGTCTGAAACCGCTCCGTCGCGCGCAC 2508  
Db |||||  
2461 GCTGCGGTATGCAACCCGAGGGTATAGGTGCTTGTCTGAAACCGCTCCGTCGCGCGCAC 2520  
Qy |||||  
2509 CTAGGTTTCGGCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGG 2568  
Db |||||  
2521 CTAGGTTTCGGCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGG 2580  
Qy |||||  
2569 GTAAGGACCATCAACCGGTGCGGCCCATCAACGCTTCTCCACCTATGGAAGGTTTCTTGGC 2628



Db 2581 GTAAGACCATCACACAGGGTGCCTCCCATCACTACTCCACCTATGCGAAGTTTCTTGCC 2640  
QY 2629 GACGFTGTTGCTCTGGGGCGCTTATGACATCATATAATATGATGAGTGCCACTCAACT 2688  
Db 2641 GACGFTGTTGCTCTGGGGCGCTTATGACATCATATAATATGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGAGAGCGGTGAGCG 2748  
Db 2701 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGAGAGCGGTGAGCG 2760  
QY 2749 CGACTCGTCTGCTGCCACCGCTACGCTCGGGATCGGTACCTGTCACATCCACATCAAAAC 2808  
Db 2761 CGACTCGTCTGCTGCCACCGCTACGCTCGGGATCGGTACCTGTCACATCCACATCAAAAC 2820  
QY 2809 ATCGAGAGGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2868  
Db 2821 ATCGAGAGGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCCATTCRAAGAGAAATGTGAT 2928  
Db 2881 ATCGAGACCATCAAGGGGGAGGACCTCATTTTCTGCCATTCRAAGAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGCGGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCCTT 2988  
Db 2941 GAGCTCGCGCGGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCCTT 3000  
QY 2989 GATGTATCCGTANTACCAACTAGCGGAGAGCTCAATTTGTGTAGCAACGAGACGCTCTAATG 3048  
Db 3001 GATGTATCCGTANTACCAACTAGCGGAGAGCTCAATTTGTGTAGCAACGAGACGCTCTAATG 3060  
QY 3049 ACGGCTTTTACGGCGATTTGACTGAGTATGACTGCAATACATGTGTCAACCGAGACA 3108  
Db 3061 ACGGCTTTTACGGCGAGCTTGCAGTCACTGAGTCAATGCTGCAATACATGTGTCAACCGAGACA 3120  
QY 3109 GTGCACTTCAGCCTGAGCCGACCTTCAACATTTAGACACAGACCGCTGCCAAGACGCG 3168  
Db 3121 GTGCACTTCAGCCTGAGCCGACCTTCAACATTTAGACACAGACCGCTGCCAAGACGCG 3180  
QY 3169 GTGTACGCTCGAGCGGGAGGACAGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3228  
Db 3181 GTGTACGCTCGAGCGGGAGGACAGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3240  
QY 3229 GTGACTTCAGGAGAACGGCCCTCGGCATGTTGCAATTCCTCGTTCCTGCGAGTGTCTAT 3288  
Db 3241 GTGACTTCAGGAGAACGGCCCTCGGCATGTTGCAATTCCTCGTTCCTGCGAGTGTCTAT 3300  
QY 3289 GACGCGGGTGTCTGTGTACGAGCTCACGCCCGCCGAGACCTCAGTTAGTTTGCGGGCT 3348  
Db 3301 GACGCGGGTGTCTGTGTACGAGCTCACGCCCGCCGAGACCTCAGTTAGTTTGCGGGCT 3360  
QY 3349 TACCTAAACACACAGGGTTGCCGTCTGCGAGGACCACTGTGAGTTCTGGAGAGCGTC 3408  
Db 3361 TACCTAAACACACAGGGTTGCCGTCTGCGAGGACCACTGTGAGTTCTGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3468  
Db 3421 TTTACAGGCTCACCCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3480  
QY 3469 AACTTCCCTTACCTGTGTAGCATACAGGCTACCGGTGTGCCAGGGCTCAGGCTCCACCT 3528  
Db 3481 AACTTCCCTTACCTGTGTAGCATACAGGCTACCGGTGTGCCAGGGCTCAGGCTCCACCT 3540  
QY 3529 CATCTGGGACCAATGTGGAAAGTGTCTCATACGGCTAAAGCCTACGCTGCAACGGGCCA 3588  
Db 3541 CATCTGGGACCAATGTGGAAAGTGTCTCATACGGCTAAAGCCTACGCTGCAACGGGCCA 3600  
QY 3589 ACGCCCTCTGTATAGGCTGGGAGCGGTTCAAAACGAGTTACTACCACACACCCCAT 3648  
Db 3601 ACGCCCTCTGTATAGGCTGGGAGCGGTTCAAAACGAGTTACTACCACACACCCCAT 3660  
QY 3649 ACCAAATACATCATGATCATGTGGGTGACCTGAGGTGCTGACGAGACCTGGGTG 3708  
Db 3661 ACCAAATACATCATGATCATGTGGGTGACCTGAGGTGCTGACGAGACCTGGGTG 3720

QY 3709 CTGTAGGCGGAGTCTTAGCAGCTCTGGCGCGGTATTTGCTGTGACAAACAGACAGCGTGGTC 3768  
Db 3721 CTGTAGGCGGAGTCTTAGCAGCTCTGGCGCGGTATTTGCTGTGACAAACAGACAGCGTGGTC 3780  
QY 3769 ATTGTGGGAGGATCATCTTGTCCGGAAGCCGGCCATCATTTCCGACAGGGAAGTCCCT 3828  
Db 3781 ATTGTGGGAGGATCATCTTGTCCGGAAGCCGGCCATCATTTCCGACAGGGAAGTCCCT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGGCCTCACACCTCCCTTACATCGAACAGGGA 3888  
Db 3841 TACCGGAGTTCGATGAGATGGAAGTGGCCTCACACCTCCCTTACATCGAACAGGGA 3900  
QY 3889 ATGAGCTCGCGCAACAATTTCAAACAGAAAGCAATCGGGTGTGTGCAACAGCACCAAG 3948  
Db 3901 ATGAGCTCGCGCAACAATTTCAAACAGAAAGCAATCGGGTGTGTGCAACAGCACCAAG 3960  
QY 3949 CAAGCGAGGCTGCTCTCCCGTGTGGAATCAAAGTGGGACCCCTCGAAGCCTTCTGG 4008  
Db 3961 CAAGCGAGGCTGCTCTCCCGTGTGGAATCAAAGTGGGACCCCTCGAAGCCTTCTGG 4020  
QY 4009 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTGTGCAACCCCGGATAGCATCACTGATGGGATTCACAGCCTTATACAGCCCGCTC 4128  
Db 4081 CTGTGCAACCCCGGATAGCATCACTGATGGGATTCACAGCCTTATACAGCCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTTAACTCTGGGGGATGGTGGCCGCCCAACTTGT 4188  
Db 4141 ACCACCAACATACCTCTCTTTTAACTCTGGGGGATGGTGGCCGCCCAACTTGT 4200  
QY 4189 CCTCCAGCGCTCTCTCTCTTTAGCGCGGATCGCTGGAGCGGTGTGTGGCAGC 4248  
Db 4201 CCTCCAGCGCTCTCTCTCTTTAGCGCGGATCGCTGGAGCGGTGTGTGGCAGC 4260  
QY 4249 ATAGGCTTTGGAAAGTGTGTGTGATTTTGGCAGGTTATGGACAGGGTGGCAGC 4308  
Db 4261 ATAGGCTTTGGAAAGTGTGTGTGATTTTGGCAGGTTATGGACAGGGTGGCAGC 4320  
QY 4309 GCGCTCGTGGCTTTAAGGTCTAGCGCGGAGATGCCCTCCACCGAGACCTGGTTAAC 4368  
Db 4321 GCGCTCGTGGCTTTAAGGTCTAGCGCGGAGTTGCCCTCCACCGAGACCTGGTTAAC 4380  
QY 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGCGCAGCGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGCGCAGCGATA 4440  
QY 4429 CTGCGTGGGACGCTGGGCGGAGGGGGTGTGCACTGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGCGTGGGACGCTGGGCGGAGGGGGTGTGCACTGATGAACCGGCTGATAGCG 4500  
QY 4489 TTCGCTTCGCGGGTAAACACAGTCTCCCGCAGCCTATGTGCTGAGAGCGAGCTGCA 4548  
Db 4501 TTCGCTTCGCGGGTAAACACAGTCTCCCGCAGCCTATGTGCTGAGAGCGAGCTGCA 4560  
QY 4549 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTAC 4608  
Db 4561 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTAC 4620  
QY 4609 CAGTGTGATCAACAGGAGTGTCTCAGCCCATGCTCGGCTCGTGGCTAAGAGATGTTGG 4668  
Db 4621 CAGTGTGATCAACAGGAGTGTCTCAGCCATGCTCGGCTCGTGGCTAAGAGATGTTGG 4680  
QY 4669 GATTGTGATGACAGGTTGATGATTTCAAGACTGGCTCCAGTCCAGTCCAGTCCCTGCCG 4728  
Db 4681 GATTGTGATGACAGGTTGATGATTTCAAGACTGGCTCCAGTCCAGTCCCTGCCG 4740  
QY 4729 CGATTCCCGGGAGTCCCTCTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGC 4788  
Db 4741 CGATTCCCGGGAGTCCCTCTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGC 4800



6961	ATCTGTGAAGCGGGGACCCAAAGAGGACGAGCGGAGCCTACGGGCTTTCACGGAGGCT	7020
7009	ATGACTAGATACTCTGCCCCCTCTGGGGACCCGCCCCAAACAGAAATACGACTTGGAGTTG	7068
7021	ATGACTAGATACTCTGCCCCCTCTGGGGACCCGCCCCAAACAGAAATACGACTTGGAGTTG	7080
7069	ATAACATCATGCTCTCCAATGTGTAGTTCGGCAGCATGTCATCTGSCAAAAGSGTGTAC	7128
7081	ATAACATCATGCTCTCCAATGTGTAGTTCGGCAGCATGTCATCTGSCAAAAGSGTGTAC	7140
7129	TATCTCACCGGTGACCCACACACCCCTTTGGCGGGCTCGCTGGGAGACAGCTAGACAC	7188
7141	TATCTCACCGGTGACCCACACACCCCTTTGGCGGGCTCGCTGGGAGACAGCTAGACAC	7200
7189	ACTCCAGTCAATTCTCTGGCTAGGCAACATCATCATGTATGCGGCCACCTTGTGGGCAAGG	7248
7201	ACTCCAGTCAATTCTCTGGCTAGGCAACATCATCATGTATGCGGCCACCTTGTGGGCAAGG	7260
7249	ATGATCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACATTGAAAAGCC	7308
7261	ATGATCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACATTGAAAAGCC	7320
7309	CTAGATGTGCAGATCTACGGGGCTGTACTTCCATTGAGCCACTGTACCTTACCTCAGATC	7368
7321	CTAGATGTGCAGATCTACGGGGCTGTACTTCCATTGAGCCACTGTACCTTACCTCAGATC	7380
7369	ATTTCACGACTCCATGCGCTTAGGGCAATTTTCACTCCATAGTTACTCTCAGGTGAGATC	7428
7381	ATTTCACGACTCCATGCGCTTAGGGCAATTTTCACTCCATAGTTACTCTCAGGTGAGATC	7440
7429	AATAGGTGGCTTTCATGCCCTCAGAAACTTGGGGTACCGGCTTGGCAGTCTCGAGACAT	7488
7441	AATAGGTGGCTTTCATGCCCTCAGAAACTTGGGGTACCGGCTTGGCAGTCTCGAGACAT	7500
7489	CGGGCCAGAAGTGTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCATTGTGGC	7548
7501	CGGGCCAGAAGTGTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCATTGTGGC	7560
7549	AAGTACCTCTTCAACTGGGCAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG	7608
7561	AAGTACCTCTTCAACTGGGCAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG	7620
7609	TCCAGTGTGATTATTCACGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAC	7668
7621	TCCAGTGTGATTATTCACGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAC	7680
7669	AGCGTGTCTGTGCGGACCGCGCTGTTTCATGTGTGGCTACTCTCTACTTTCTGTAGGG	7728
7681	AGCGTGTCTGTGCGGACCGCGCTGTTTCATGTGTGGCTACTCTCTACTTTCTGTAGGG	7740
7729	GTAGGCATCTATCTACTCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGG	7788
7741	GTAGGCATCTATCTACTCCCAACCGATGAACGGGGAGCTAAACACTCCAGGCCAATAGG	7800
7789	CCATCTGTTTTTTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	7848
7801	CCATCTGTTTTTTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	7860
7849	TTTTCTCTTTTTTTTTTTCCTTTTTTTTTTTCCTTTTTCTTTCTTTGGTGGCTCAATCTAGC	7908
7861	TTTTCTCTTTTTTTTTTTCCTTTTTTTTTTTCCTTTTTCTTTCTTTGGTGGCTCAATCTAGC	7920
7909	CCTAGTCACGGCTAGCTGTGAAAAGTCCGTCGAGCCGCTTGACTTCGACGAGAGTCTGTATC	7968
7921	CCTAGTCACGGCTAGCTGTGAAAAGTCCGTCGAGCCGCTTGACTTCGACGAGAGTCTGTATC	7980
7969	TGGCCTCTCTGCAGATCAAGT	7989
7981	TGGCCTCTCTGCAGATCAAGT	8001

RESULT 13  
AX036261

QY 709 CTCCTCCGAGAAAGTATCCATCATGCTGATGCAATCGGCGGCTGCATACGCTTGATC 768  
DB 721 CTCCTGCCGAGAAAGTATCCATCATGCTGATGCAATCGGCGGCTGCATACGCTTGATC 780  
QY 769 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGCTACTCGGA 828  
DB 781 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGCTACTCGGA 840  
QY 829 TGGAAAGCGGCTTCTGATCAGGATGATCTGGAAGAGAGATCAGGGGCTCGCGCAG 888  
DB 841 TGGAAAGCGGCTTCTGATCAGGATGATCTGGAAGAGAGATCAGGGGCTCGCGCAG 900  
QY 889 CGGAAGCTTCGCGAGGCTCAAGCGCGCATCGCGAGCGAGATCTCGTCTGACCC 948  
DB 901 CGGAAGCTTCGCGAGGCTCAAGCGCGCATCGCGAGCGAGATCTCGTCTGACCC 960  
QY 949 ATGGCGATGCTGCTTGGCGAATATCATGTTGGAATAATGGCGCTTTCTGGATTCATCG 1008  
DB 961 ATGGCGATGCTGCTTGGCGAATATCATGTTGGAATAATGGCGCTTTCTGGATTCATCG 1020  
QY 1009 ACTGTGGCGGCTGGGTGTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCGTGATA 1068  
DB 1021 ACTGTGGCGGCTGGGTGTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCGTGATA 1080  
QY 1069 TTGCTGAAGAGCTTCGCGCGAATCGGCTGACCGCTTCTGCTGCTTTACGCTATCGCG 1128  
DB 1081 TTGCTGAAGAGCTTCGCGCGAATGGGCTGACCGCTTCTGCTGCTTTACGCTATCGCG 1140  
QY 1129 CTCGCGATTTCGAGCGCATCGCTTCTATCGCTTCTTTCGAGTTCCTCTGAGTTTAAA 1188  
DB 1141 CTCGCGATTTCGAGCGCATCGCTTCTATCGCTTCTTTCGAGTTCCTCTGAGTTTAAA 1200  
QY 1189 CAGACCAACAGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTT 1248  
DB 1201 CAGACCAACAGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTT 1260  
QY 1249 AACGTTACTGGCGAAGCGCTTGAATAAGCGCGGTGTGCGTTTGTCTATATGTTATT 1308  
DB 1261 AACGTTACTGGCGAAGCGCTTGAATAAGCGCGGTGTGCGTTTGTCTATATGTTATT 1320  
QY 1309 TCCACCATATTCGCTTCTTTGGCAATGTAGGGCCCGGAAACCTTGGCCCTGCTCTTTG 1368  
DB 1321 TCCACCATATTCGCTTCTTTGGCAATGTAGGGCCCGGAAACCTTGGCCCTGCTCTTTG 1380  
QY 1369 ACGACATTCCTAGGGGTCTTTCCTCTCGCAAGGAAATGCAAGGTCTGTGTAATGTC 1428  
DB 1381 ACGACATTCCTAGGGGTCTTTCCTCTCGCAAGGAAATGCAAGGTCTGTGTAATGTC 1440  
QY 1429 GTGAAGGAGCAGTTTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCTT 1488  
DB 1441 GTGAAGGAGCAGTTTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCTT 1500  
QY 1489 TGCAGGACGGAACCCCGCACTCGGCAAGGTGCTCTCGGCGCAAAAGCCACGTGTA 1548  
DB 1501 TGCAGGACGGAACCCCGCACTCGGCAAGGTGCTCTCGGCGCAAAAGCCACGTGTA 1560  
QY 1549 TAAGATACCTGCAAGGCGGCAACCCAGTCCAGTTCAGTTGTAGTTGATGTTG 1608  
DB 1561 TAAGATACCTGCAAGGCGGCAACCCAGTCCAGTTCAGTTGTAGTTGATGTTG 1620  
QY 1609 GAAAGGTCAATGGCTCTCTCTCAAGCGTATTCAAGAGGGCTGAAGATGCCAGAG 1668  
DB 1621 GAAAGGTCAATGGCTCTCTCTCAAGCGTATTCAAGAGGGCTGAAGATGCCAGAG 1680  
QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTTTAG 1728  
DB 1681 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTTTAG 1740  
QY 1729 TCGAGGTTAAAAAAGCTCTAGGCCCGCCGGAACGAGGAGCTGTTTCTTTTCAAAAA 1788  
DB 1741 TCGAGGTTAAAAAAGCTCTAGGCCCGCCGGAACGAGGAGCTGTTTCTTTTCAAAAA 1800  
QY 1789 CACGATAATACCATGGCGCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGC 1848

DB 1801 CACGATAATACCATGGCGCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGC 1860  
QY 1849 TGCATCATCTAGCTCCTCAGAGCGGACAGGAACCAAGGTGAGGGGAGGTCCAAAGTG 1908  
DB 1861 TGCATCATCTAGCTCCTCAGAGCGGACAGGAACCAAGGTGAGGGGAGGTCCAAAGTG 1920  
QY 1909 GTCTCCACGGCAACAAATCTTTCTTGGGACCTGCGTCAATGGGCTGTGTGGAATGTC 1968  
DB 1921 GTCTCCACGGCAACAAATCTTTCTTGGGACCTGCGTCAATGGGCTGTGTGGAATGTC 1980  
QY 1969 TATCATGTGCGGCTCAAGAACCTTCCCGCCCAAGGGCCCAATCACCCAAATGTAC 2028  
DB 1981 TATCATGTGCGGCTCAAGAACCTTCCCGCCCAAGGGCCCAATCACCCAAATGTAC 2040  
QY 2029 ACCAATGTGACAGGACCTCTGCTGGCTGGCAAGCCCGCCCGGGCGCTTCTTGTACA 2088  
DB 2041 ACCAATGTGACAGGACCTCTGCTGGCTGGCAAGCCCGCCCGGGCGCTTCTTGTACA 2100  
QY 2089 CCATGCACCTGGGCGAGCTCGGACCTTTTACCTTGGTTCAGAGGCATGCGATGTCATTCG 2148  
DB 2101 CCATGCACCTGGGCGAGCTCGGACCTTTTACCTTGGTTCAGAGGCATGCGATGTCATTCG 2160  
QY 2149 GTGCGCCCGGCGGACAGAGGGGAGCCTACTCTCCCGCAGGCCCTCTCTCTACTTGG 2208  
DB 2161 GTGCGCCCGGCGGACAGAGGGGAGCCTACTCTCTCCCGCAGGCCCTCTCTCTACTTGG 2220  
QY 2209 AAGGCTCTTTCGGGCGGTCCTCTGCTGCGCCCTCGGGGACGCTGTGGGCAATCTTTCG 2268  
DB 2221 AAGGCTCTTTCGGGCGGTCCTCTGCTGCGCCCTCGGGGACGCTGTGGGCAATCTTTCG 2280  
QY 2269 GCTGCGGCTGTCACCCGAGGGGTTCCGAAGGGGTTGACTTTGTACCGTTCGAGTCTATG 2328  
DB 2281 GCTGCGGCTGTCACCCGAGGGGTTCCGAAGGGGTTGACTTTGTACCGTTCGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGCTCCCGGCTCTTACGGACAACTCGTCCCTTCGCGCGCTACCGAG 2388  
DB 2341 GGAACCACTATGCGGCTCCCGGCTCTTACGGACAACTCGTCCCTTCGCGCGCTACCGAG 2400  
QY 2389 ACATTCCAGGTGGCCATCTACAGCCCTACTGTAGCGGCAAGACACTAAGTGTGCG 2448  
DB 2401 ACATTCCAGGTGGCCATCTACAGCCCTACTGTAGCGGCAAGACACTAAGTGTGCG 2460  
QY 2449 GCTGGTATGAGCCCAAGGTTAAGTGTCTTCTGCTGAACCGCTCGTTCGCGCCGACC 2508  
DB 2461 GCTGGTATGAGCCCAAGGTTAAGTGTCTTCTGCTGAACCGCTCGTTCGCGCCGACC 2520  
QY 2509 CTAGGTTTCGGGCGGTATATGTCTAAGGCAATGGTATCGACCTTAAACATCAGAAACCG 2568  
DB 2521 CTAGGTTTCGGGCGGTATATGTCTAAGGCAATGGTATCGACCTTAAACATCAGAAACCG 2580  
QY 2569 GTAAGGACCATCAACCGGGTCCCGCATACGTAATCGAATGGAAGGTTCTTGGC 2628  
DB 2581 GTAAGGACCATCAACCGGGTCCCGCATACGTAATCGAATGGAAGGTTCTTGGC 2640  
QY 2629 GACGTTGTTGCTCTGGGGCGCTTACACATCAATATGATGATGAGTGCACCTCAACT 2688  
DB 2641 GACGTTGTTGCTCTGGGGCGCTTACACATCAATATGATGATGAGTGCACCTCAACT 2700  
QY 2689 GACTCGACCATATCTCTGGGCGATCGGCACAGTCTCTGACCAAGCGGAGCGCTGGAGCG 2748  
DB 2701 GACTCGACCATATCTCTGGGCGATCGGCACAGTCTCTGACCAAGCGGAGCGCTGGAGCG 2760  
QY 2749 CGACTGCTGCTGCTCGCACCGCTACGCTCGGGATCGGTACCGTGCACATCAAAAC 2808  
DB 2761 CGACTGCTGCTGCTCGCACCGCTACGCTCGGGATCGGTACCGTGCACATCAAAAC 2820  
QY 2809 ATCGAGGAGTGGCTCTCTCCAGCACTCGAATAATCCCTTTTATGCGAAAGCCATCCCC 2868  
DB 2821 ATCGAGGAGTGGCTCTCTCCAGCACTCGAATAATCCCTTTTATGCGAAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGGAGGCACTCATTTTCTGCCATTCCTCAAGAAATGTGAT 2928

Db 2881 ATCGAGACCATCAAGGGGGAGCGACCTCATTTTCTGCCATTTCCAGAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGCGAAGCTGTCCGSCCTCGGACTCAATGCTGTAGCAATATACCGGGGCCCTT 2988  
Db 2941 GAGCTCGCGCGAAGCTGTCCGSCCTCGGACTCAATGCTGTAGCAATATACCGGGGCCCTT 3000  
QY 2989 GATGATCCGTATACCAATAGCGAGAGCGTCAATGCTGTAGCAAGCGAGCTCTAAATG 3048  
Db 3001 GATGATCCGTATACCAATAGCGAGAGCGTCAATGCTGTAGCAAGCGAGCTCTAAATG 3060  
QY 3049 ACGGGCTTTACCGCGCATTTCCGACTCAGTGATCGAATGCAATACATGTGTCAACCCAGACA 3108  
Db 3061 ACGGGCTTTACCGCGCATTTCCGACTCAGTGATCGAATGCAATACATGTGTCAACCCAGACA 3120  
QY 3109 GTCGACTTCAGCTGAGCCGCACTTCAACANTGAGACGACGACCGTGGCCACAAGACGG 3168  
Db 3121 GTCGACTTCAGCTGAGCCGCACTTCAACANTGAGACGACGACCGTGGCCACAAGACGG 3180  
QY 3169 GTGTACGCTCGCAGCGGCGAGGACGACTGGTAGGGGCGAGATGGGCATTTACAGTTT 3228  
Db 3181 GTGTACGCTCGCAGCGGCGAGGACGACTGGTAGGGGCGAGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTTCCAGGAAACCGGCCCTCGGCGATGTTCCGATTCCTCGGTTCTGTGCGAGTGCTAT 3288  
Db 3241 GTGACTTCCAGGAAACCGGCCCTCGGCGATGTTCCGATTCCTCGGTTCTGTGCGAGTGCTAT 3300  
QY 3289 GAGCGGGCTGTCTTGTGTACGAGCTACGCCCGCGAGACCTCAGTTAGTTTGGGGCT 3348  
Db 3301 GAGCGGGCTGTCTTGTGTACGAGCTACGCCCGCGAGACCTCAGTTAGTTTGGGGCT 3360  
QY 3349 TACCTAAACACACCGAGTTGCCGCTGTCAGGACCATCTGGAGTCTTGGGAGAGCGTC 3408  
Db 3361 TACCTAAACACACCGAGTTGCCGCTGTCAGGACCATCTGGAGTCTTGGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGCGAGAGAC 3468  
Db 3421 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGCGAGAGAC 3480  
QY 3469 AACTTCCCTTACTGTGTAGATACAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCT 3528  
Db 3481 AACTTCCCTTACTGTGTAGATACAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCT 3540  
QY 3529 CCATCGTGGGACCAATGTGGAGTCTCTATACGGCTAAAGCTACGCTGACGCGGCA 3588  
Db 3541 CCATCGTGGGACCAATGTGGAGTGTCTATACGGCTAAAGCTACGCTGACGCGGCA 3600  
QY 3589 ACGCCCTGCTGTATAGGCTGGGAGCGGTTCAAAACGAGTTACTACACACACCCCAT 3648  
Db 3601 ACGCCCTGCTGTATAGGCTGGGAGCGGTTCAAAACGAGTTACTACACACACCCCAT 3660  
QY 3649 ACCAAATACATATGGCATGTGCTGGCTGACCTGGAGTGTGTACGAGACCTGGGTG 3708  
Db 3661 ACCAAATACATATGGCATGTGCTGGCTGACCTGGAGTGTGTGTACGAGACCTGGGTG 3720  
QY 3709 CTGTAGGCGGAGTCTTACGAGCTGTGGCGGTATTTCTGTACACAGCAGCGTGTGTC 3768  
Db 3721 CTGTAGGCGGAGTCTTACGAGCTGTGGCGGTATTTCTGTACACAGCAGCGTGTGTC 3780  
QY 3769 ATTGTGGGAGGATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAATTCCTT 3828  
Db 3781 ATTGTGGGAGGATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAATTCCTT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGTGCTCACACTCCCTTACATTCGAACAGGGA 3888  
Db 3841 TACCGGAGTTCGATGAGATGGAAGTGTGCTCACACTCCCTTACATTCGAACAGGGA 3900  
QY 3889 ATGACGCTCCGAAACAAATTAACAGAGGCAATCGGTTGTGTGCAACAGCCACAAG 3948  
Db 3901 ATGACGCTCCGAAACAAATTAACAGAGGCAATCGGTTGTGTGCAACAGCCACAAG 3960  
QY 3949 CAAGCGGAGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCTCGAAGCTTCTGG 4008  
Db 3961 CAAGCGGAGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCTCGAAGCTTCTGG 4020

QY 4009 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATAATTTAGCAGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATAATTTAGCAGCTTGTCCACTCTG 4080  
QY 4069 CTGGCAACCCCGGATAGCATCACTGATGGGATTCACAGCTCTATCACCAGCCGCTC 4128  
Db 4081 CTGGCAACCCCGGATAGCATCACTGATGGGATTCACAGCTCTATCACCAGCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTTAAATCCTGGGGGATGGGTGGCCGCCCAACTTGTCT 4188  
Db 4141 ACCACCAACATACCTCTCTTTTAAATCCTGGGGGATGGGTGGCCGCCCAACTTGTCT 4200  
QY 4189 CTTCCAGCGCTGTCTCTCTCTTGTGAGCGCGGCAATCGCTGGACGCGTGTGGGAGC 4248  
Db 4201 CTTCCAGCGCTGTCTCTCTCTCTTGTGAGCGCGGCAATCGCTGGAGCGGCTGTGGCAGC 4260  
QY 4249 ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGAGGTTATGGAGCAGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGAGGTTATGGAGCAGGGTGGCAGGC 4320  
QY 4309 GCGCTCGTGGCTTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAAC 4368  
Db 4321 GCGCTCGTGGCTTTAAGGTCATGAGCGCGAGTGGCCCTCCACCGAGGACCTGGTTAAC 4380  
QY 4369 CTACTCCCTGTCTATCT 4428  
Db 4381 CTACTCCCTGTCTATCT 4440  
QY 4429 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTCTCAGTGGATGAACCGCTGATAGCG 4488  
Db 4441 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTCTCAGTGGATGAACCGCTGATAGCG 4500  
QY 4489 TTCGCTTCCGCGGTAAACACAGCTCTCCCCACGCACTATGTGCTCAGAGCGCAGCTGCA 4548  
Db 4501 TTCGCTTCCGCGGTAAACACAGCTCTCCCCACGCACTATGTGCTCAGAGCGCAGCTGCA 4560  
QY 4549 GCACGTGTCACTCAGATCT 4608  
Db 4561 GCACGTGTCACTCAGATCT 4620  
QY 4609 CAGTGGATCAACGAGGACTGCTCCAGCCATGCTCCGCGCTCGTGGCTAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACGAGGACTGCTCCAGCCATGCTCCGCGCTCGTGGCTAGAGATGTTGG 4680  
QY 4669 GATTGGATATGACGCGTGTGACTGATTTCAAGACCTCGCTCCAGTCCCAAGCTCTCTGCCG 4728  
Db 4681 GATTGGATATGACGCGTGTGACTGATTTCAAGACCTCGCTCCAGTCCCAAGCTCTCTGCCG 4740  
QY 4729 CGATTGCGGGAGTCCCT 4788  
Db 4741 CGATTGCGGGAGTCCCT 4800  
QY 4789 GACGCGATCATCAACACACCTGCGCATCTGGAGCACAGATCACCGGACATGTGAAAAAC 4848  
Db 4801 GACGCGATCATCAACACACCTGCGCATCTGGAGCACAGATCACCGGACATGTGAAAAAC 4860  
QY 4849 GGTTCATGAGGATCGTGGGCTCTAGGACCTGTAGTAAACACGTGGCATGGAACATTTCCCT 4908  
Db 4861 GGTTCATGAGGATCGTGGGCTCTAGGACCTGTAGTAAACACGTGGCATGGAACATTTCCCT 4920  
QY 4909 ATTACCGGTAACACCGGCGCTCTGACGCGCTCCCGCGGCCAAATTTCTAGGGCG 4968  
Db 4921 ATTACCGGTAACACCGGCGCTCTGACGCGCTCCCGCGGCCAAATTTCTAGGGCG 4980  
QY 4969 CTGTGGCGGTGCTCTGAGGAGTACGTGGAGTTACCGGGTGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGTGCTCTGAGGAGTACGTGGAGTTACCGGGTGGGATTTCCACTAC 5040  
QY 5029 GTGAGGGGATGACCACTGACAACTGAAAGTGGCGGTGTGAGGTTCCGCGGCCCGGAATTC 5088  
Db 5041 GTGAGGGGATGACCACTGACAACTGAAAGTGGCGGTGTGAGGTTCCGCGGCCCGGAATTC 5100

QY 5089 TTCAAGAGTGGATGGGGTGGGTTGCAAGGTACGCTCCAGCGTGCAGAACCCCTCCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGGTTGCAAGGTACGCTCCAGCGTGCAGAACCCCTCCTA 5160  
QY 5149 CGGAGGAGGTACATTCCTGCTGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCA 5208  
Db 5161 CGGAGGAGGTACATTCCTGCTGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCA 5220  
QY 5209 TCGAGCCGCAACCGGAGTGTAGCTGCTCACTTCATGCTCACCGACCCCTCCACATTT 5268  
Db 5221 TCGAGCCGCAACCGGAGTGTAGCTGCTCACTTCATGCTCACCGACCCCTCCACATTT 5280  
QY 5269 ACGGCGGAGCGGTAAAGCTAGGCTGGCAGGGGATCTCCGCCCTCTTGGCCAGCTCA 5328  
Db 5281 ACGGCGGAGCGGTAAAGCTAGGCTGGCAGGGGATCTCCGCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGTACGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGTCATGACTCC 5388  
Db 5341 TCAGTACGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGTCATGACTCC 5400  
QY 5389 CGGACCGTGAACCTCATCGAGCCCAACCTCTGTGGCGGAGGATGGCGGGAGACATC 5448  
Db 5401 CGGACCGTGAACCTCATCGAGCCCAACCTCTGTGGCGGAGGATGGCGGGAGACATC 5460  
QY 5449 ACCCGCGTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGC 5508  
Db 5461 ACCCGCGTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGC 5520  
QY 5509 GAGGAGATGAGAGGGAATGTCGTTCCGGCGGAGATCTTTCGCGAGGTCCAGGAAATTC 5568  
Db 5521 GAGGAGATGAGAGGGAATGTCGTTCCGGCGGAGATCTTTCGCGAGGTCCAGGAAATTC 5580  
QY 5569 CCTCGAGGATGCCATATGGCGAGCGCGGATACAACTCCACTGTAGAGTCTCTGG 5628  
Db 5581 CCTCGAGGATGCCATATGGCGAGCGCGGATACAACTCCACTGTAGAGTCTCTGG 5640  
QY 5629 AAGGACCCGGACTACGTCCTCCAGTGGTACACGGGTGTCAATTCGCCCTGCCAAGGCC 5688  
Db 5641 AAGGACCCGGACTACGTCCTCCAGTGGTACACGGGTGTCAATTCGCCCTGCCAAGGCC 5700  
QY 5689 CCTCGATACCACTTCCAAGGAGGAGGACGTTGTCTGTCAAGATCTACCGTGTCT 5748  
Db 5701 CCTCGATACCACTTCCAAGGAGGAGGACGTTGTCTGTCAAGATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGCGAGCTCGCCACAAAGACCTTCGACGCTCGGATCGTGGCGCTCGAC 5808  
Db 5761 TCTGCTTGGCGAGCTCGCCACAGAGACCTTCGCGACGCTCGGAATCGTGGCGCTCGAC 5820  
QY 5809 AGCGGACGGCAACGGCCCTCTCCTGACAGGCCCTCCGACGCGCGGATCCGAC 5868  
Db 5821 AGCGGACGGCAACGGCCCTCTCCTGACAGGCCCTCCGACGCGCGGATCCGAC 5880  
QY 5869 GTTGAATGCTACTCTCTCATGCCCTCTTGAAGGAGGAGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAATGCTACTCTCTCATGCCCTCTTGAAGGAGGAGGATCCCGATCTCAGC 5940  
QY 5929 GACGGCTCTGCTTACCGTAAAGGAGGAGGCTAGTAGAGGACGTCGTCGTCGTCGATG 5988  
Db 5941 GACGGCTCTGCTTACCGTAAAGGAGGAGGCTAGTAGAGGACGTCGTCGTCGTCGATG 6000  
QY 5989 TCCTACATGAGCAGGCGCCCTGATACGCCATGCGCTGGCGGAGAAACCAAGTGGCC 6048  
Db 6001 TCCTACATGAGCAGGCGCCCTGATACGCCATGCGCTGGCGGAGAAACCAAGTGGCC 6060  
QY 6049 ATCAATGACCTGAGCACTCTTGTGCTCGGTACCAACCACTTGTGTATGTCTACACATCT 6108  
Db 6061 ATCAATGACCTGAGCACTCTTGTGCTCGGTACCAACCACTTGTGTATGTCTACACATCT 6120  
QY 6109 CGCAGCGCAAGCTGCGCAGAGAGGTCACTTTTGACAGACTGACAGTCTCTGAGCAGC 6168  
Db 6121 CGCAGCGCAAGCTGCGCAGAGAGGTCACTTTTGACAGACTGACAGTCTCTGAGCAGC 6180  
QY 6169 CACTACCGGAGCGTGTCTCAAGAGATGAAGCGGATCCACAGTTAAGGCTAAACTT 6228

Db 6181 CACTACCGGAGCGTGTCTCAAGAGATGAAGCGGAGGCTCCACAGTTAAGCTAAACTT 6240  
QY 6229 CTATCCGTGGAGGAGCGCTGTAAAGCTGACGCCCCACATTCGGCGAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGGAGGAGCGCTGTAAAGCTGACGCCCCACATTCGGCGAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAAAGGAGGCTCGGAACTTATCCAGGAGGCGGTTTACCAATCGGCTCCGTCG 6348  
Db 6301 TATGGGCAAAAGGAGGCTCGGAACTTATCCAGGAGGCGGTTTAAACCATCGGCTCCGTCG 6360  
QY 6349 TGGAGGACTTGTCTGGAGACACTGAGACACCAATTGACACCACTCATGGCAAAAAT 6408  
Db 6361 TGGAGGACTTGTCTGGAGACACTGAGACACCAATTGACACCACTCATGGCAAAAAT 6420  
QY 6409 GAGGTTTTCGCTCCAAACAGAGAGGCGGCGCAAGCGAGCTCGCTTATCGTATTC 6468  
Db 6421 GAGGTTTTCGCTCCAAACAGAGAGGCGGCGCAAGCGAGCTCGCTTATCGTATTC 6480  
QY 6469 CCAGATTTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTC 6528  
Db 6481 CCAGATTTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTC 6540  
QY 6529 CCTCAGGCGGTGATGGGCTCTTTCATACGGAATTCCTTCTCTGGACAGCGGTTCGAG 6588  
Db 6541 CCTCAGGCGGTGATGGGCTCTTTCATACGGAATTCCTTCTCTGGACAGCGGTTCGAG 6600  
QY 6589 TTCTGTGTGAATGCTGTGAAAGCGAAATGCGCTTATGGGCTTCGCAATGACACCGC 6648  
Db 6601 TTCTGTGTGAATGCTGTGAAAGCGAAATGCGCTTATGGGCTTCGCAATGACACCGC 6660  
QY 6649 TGTTTTGACTCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6708  
Db 6661 TGTTTTGACTCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAAGTCTGCTCACAGACGCGCTTTACATC 6768  
Db 6721 TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAAGTCTGCTCACAGACGCGCTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6828  
Db 6781 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6840  
QY 6829 GGTGTACTGACGACGAGCTGGGTAATACCTCATCTGTTACTTTGAAGCGCGCTCGGCC 6888  
Db 6841 GGTGTACTGACGACGAGCTGGGTAATACCTCATCTGTTACTTTGAAGCGCGCTCGGCC 6900  
QY 6889 TGTGAGCTGCGAAGCTCCAGGACTGCAACGATGCTCTGATGCGGAGACGACCTTCTCGTT 6948  
Db 6901 TGTGAGCTGCGAAGCTCCAGGACTGCAACGATGCTCTGATGCGGAGACGACCTTCTCGTT 6960  
QY 6949 ATCTGTGAAAGCGGGGAGCCCAAGAGACGAGGGAGCCTACGGGCTTACGGAGGCT 7008  
Db 6961 ATCTGTGAAAGCGGGGAGCCCAAGAGACGAGGGAGCCTACGGGCTTACGGAGGCT 7020  
QY 7009 ATGACTAGATACTCTGCCCCCTGGGACCCGCCCAACCAAGATAGACTTGGAGTTG 7068  
Db 7021 ATGACTAGATACTCTGCCCCCTGGGACCCGCCCAACCAAGATAGACTTGGAGTTG 7080  
QY 7069 ATAAATCATGCTCTCTCCAATGTGTCACTGCTGCGCAACGATGCTTGGCAAAAGGTTGAC 7128  
Db 7081 ATAAATCATGCTCTCTCCAATGTGTCACTGCTGCGCAACGATGCTTGGCAAAAGGTTGAC 7140  
QY 7129 TATCTACCCGCTGACCCCAACCCCTTGGCGGGTGTGCTGGAGACGACTAGACAC 7188  
Db 7141 TATCTACCCGCTGACCCCAACCCCTTGGCGGGTGTGCTGGAGACGACTAGACAC 7200  
QY 7189 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCTATGTATGCGCCCACTTGTGGCAAGG 7248  
Db 7201 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCTATGTATGCGCCCACTTGTGGCAAGG 7260  
QY 7249 ATGATCTGTATGACTCATTTCTCTCATCTCTAGCTCAGGAACTTGAAGGCTTAAAGGCC 7308



Db 7261 ATGATCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACCTTGAAAGGCC 7320

QY 7309 CTAGATTGTGAGATCTACGGGGCTGTACTCCANTGAGCACTTGACCTACCTCAGATC 7368

Db 7321 CTAGATTGTGAGATCTACGGGGCTGTACTCCANTGAGCACTTGACCTACCTCAGATC 7380

QY 7369 ATTCAACGACTCCATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7428

Db 7381 ATTCAACGACTCCATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7440

QY 7429 AATAGGTTGGCTTTCATGCTCAGAAACTTGGGTACCGCCCTTGGCAGTCTCGAAGACAT 7488

Db 7441 AATAGGTTGGCTTTCATGCTCAGAAACTTGGGTACCGCCCTTGGCAGTCTCGAAGACAT 7500

QY 7489 CGGGCCAGAAGTGTCCGCGTAGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7548

Db 7501 CGGGCCAGAAGTGTCCGCGTAGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7560

QY 7549 AAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7608

Db 7561 AAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7620

QY 7609 TCCAGTGTGATTTATCAGCTGGTTGGTTGCTGGTTACAGCGGGGAGACATATATCAC 7668

Db 7621 TCCAGTGTGATTTATCAGCTGGTTGGTTGCTGGTTACAGCGGGGAGACATATATCAC 7680

QY 7669 AGCTGTCTGTCGCGGACCGCGTGTTCATGTGTGCTACTCTCTACTCTTCTGTAGGG 7728

Db 7681 AGCTGTCTGTCGCGGACCGCGTGTTCATGTGTGCTACTCTCTACTCTTCTGTAGGG 7740

QY 7729 GTAGGCACTATCTACTCCCAACCGGATGAAGCGGACCTTAAACACTCCAGGCAATAGG 7788

Db 7741 GTAGGCACTATCTACTCCCAACCGGATGAAGCGGAGCTTAAACACTCCAGGCAATAGG 7800

QY 7789 CCATCTGTTTTTCCCTT 7848

Db 7801 CCATCTGTTTTTCCCTT 7860

QY 7849 TTTTTCCTTTTTTTTTTCCCTT 7908

Db 7861 TTTTTCCTTTTTTTTTTCCCTT 7920

QY 7909 CTTAGTCAAGCTGAGTGTGAAGGCTCCGTGAGCGCTTGACGACGAGAGTCTGTATC 7968

Db 7921 CTTAGTCAAGCTGAGTGTGAAGGCTCCGTGAGCGCTTGACGACGAGAGTCTGTATC 7980

QY 7969 TGGCCTCTCTGAGATCAAGT 7989

Db 7981 TGGCCTCTCTGAGATCAAGT 8001

RESULT 14  
AX937622

LOCUS 12305 bp DNA linear PAT 06-JAN-2004  
DEFINITION Sequence 2 from Patent WO03091439.  
ACCESSION AX937622  
VERSION AX937622.1 GI:40713673

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

REFERENCE 1  
AUTHORS Duggal, R. A., Patrick, A. K., Zhang, J. A. and Zhao, W. A.  
TITLE Reporter-selectable hepatitis c virus replicon  
JOURNAL Patent: WO 03091439-A 2 06-NOV-2003;  
PFIZER INC. (US)

FEATURES  
Location/Qualifiers  
1..12305  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="A cell line wherein the nucleic acid molecule is a  
self replicating RNA molecule."

## ORIGIN

Query Match 95.3%; Score 7610.4; DB 6; Length 12305;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7611; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 378 GGGCGGCCATGATTGAACAAGATGATTCACGCGAGTTCTCGGCGGCTTGGGTGAG 437

Db 1370 GGGCGGCCATGATTGAACAAGATGATTCACGCGAGTTCTCGGCGGCTTGGGTGAG 1429

QY 438 AGCTATTTCGGCTATGACTGGGCACACACAACTCGGCTGCTCTGATGCGCGCTGTC 497

Db 1430 AGCTATTTCGGCTATGACTGGGCACACAACTCGGCTGCTCTGATGCGCGCTGTC 1489

QY 498 CGGCTGTGACGCGAGGGGGCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTG 557

Db 1490 CGGCTGTGACGCGAGGGGGCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTG 1549

QY 558 AATGAATCGACGACGAGCGGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC 617

Db 1550 AATGAATCGACGACGAGCGGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC 1609

QY 618 GCAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAGTG 677

Db 1610 GCAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAGTG 1669

QY 678 CGGGCGAGGATCTCCTGTCACTCACTTGTCTCTCCGAGAAAGTATCCATCATGGCT 737

Db 1670 CGGGCGAGGATCTCCTGTCACTCACTTGTCTCTCCGAGAAAGTATCCATCATGGCT 1729

QY 738 GATGCAATGGCGGCTGATAGCTGTATCGGCTACCTGCCCATTCGACCAACGAGCG 797

Db 1730 GATGCAATGGCGGCTGATAGCTGTATCGGCTACCTGCCCATTCGACCAACGAGCG 1789

QY 798 AAACATCGCATCGAGGAGCAAGTACTCGATGGAAGCGGCTCTGTCGATCAGGATGAT 857

Db 1790 AAACATCGCATCGAGGAGCAAGTACTCGATGGAAGCGGCTCTGTCGATCAGGATGAT 1849

QY 858 CTGGACGAAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCGCCAGGCTCAAGGCGCGC 917

Db 1850 CTGGACGAAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCGCCAGGCTCAAGGCGCGC 1909

QY 918 ATGCCCGAGCGGAGGATCTCGTGTGACCATGGCGATGCTGCTTGGCAATATCATG 977

Db 1910 ATGCCCGAGCGGAGGATCTCGTGTGACCATGGCGATGCTGCTTGGCAATATCATG 1969

QY 978 GTGGAATAATGGCGCTTTTCTGGATTTCATGACTGTGGCGGCTGGGTGTGGCGGACCGC 1037

Db 1970 GTGGAATAATGGCGCTTTTCTGGATTTCATGACTGTGGCGGCTGGGTGTGGCGGACCGC 2029

QY 1038 TATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCAATGGGCT 1097

Db 2030 TATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCAATGGGCT 2089

QY 1098 GACCGCTTCTCTGCTTTTACCGGTATCGCGCTCCCGATTCCGAGCGCATTCGCTTCTAT 1157

Db 2090 GACCGCTTCTCTGCTTTTACCGGTATCGCGCTCCCGATTCCGAGCGCATTCGCTTCTAT 2149

QY 1158 CGCTTCTTACGAGTTCTTCTGAGTTTAAACAGACCAACGTTTCCCTCTAGCGGGA 1217

Db 2150 CGCTTCTTACGAGTTCTTCTGAGTTTAAACAGACCAACGTTTCCCTCTAGCGGGA 2209

QY 1218 TCAATTTCGCGGCTCTCCCTCCCCCCCCCTAACGTTACTGCCGGAAGCGCTTGAATA 1277

Db 2210 TCAATTTCGCGGCTCTCCCTCCCCCCCCCTAACGTTACTGCCGGAAGCGCTTGAATA 2269

QY 1278 AGCGGCTGTGGTTTGTCTATATGTTATTTTCCACCATTTTGGCGCTTTTGGCAATGT 1337

Db 2270 AGCGGCTGTGGTTTGTCTATATGTTATTTTCCACCATTTTGGCGCTTTTGGCAATGT 2329

QY 1338 GAGGCGCGGAAACCTGGCGGCTCTCTTGTAGCAGGACTTCCTAGGGCTTTTCCCTCT 1397

Db 2330 GAGGCGCGGAAACCTGGCGGCTCTCTTGTAGCAGGACTTCCTAGGGCTTTTCCCTCT 2389



QY	1398	CGCCAAAGGAATGCMAGTCTGTGTAATGTCGTGAGGAGCAGTTCTCTGGAGCTTC	1457	QY	2478	GCTTGTCTCTGAACCGCTCCGTCGCGCCACCTAGTTTTCGGGGCTATATCTCTAAGGC	2537
Db	2390	CGCCAAAGGAATGCAAGTCTGTGTAATGTCGTGAGGAGCAGTTCTCTGGAGCTTC	2449	Db	3470	GCTTGTCTCTGAACCGCTCCGTCGCGCCACCTAGTTTTCGGGGCTATATCTCTAAGGC	3529
QY	1458	TTGAAGACAAACAACTCTGTAGCAGCCTTTGACGAGCAGGACACCCGCCACCTGGCGA	1517	QY	2538	ACATGGTATGACCCCTAAACATCAGAACCGGGTAAAGAACCATCAACCGGGTGCCTCCAT	2597
Db	2450	TTGAAGACAAACAACTCTGTAGCAGCCTTTGACGAGCAGGACACCCGCCACCTGGCGA	2509	Db	3530	ACATGGTATGACCCCTAAACATCAGAACCGGGTAAAGAACCATCAACCGGGTGCCTCCAT	3589
QY	1518	CAGTGCCTCTGCGGCCAAAGCCACGCTGTATAGANTACCTGTCAAAAGGCGGCACAACC	1577	QY	2598	CACGTACTCCACCTATGCAAGTTCTTTCGCGAGGTTGCTCTGCGGGGCGCTATGA	2657
Db	2510	CAGTGCCTCTGCGGCCAAAGCCACGCTGTATAGANTACCTGTCAAAAGGCGGCACAACC	2569	Db	3590	CACGTACTCCACCTATGCAAGTTCTTTCGCGAGGTTGCTCTGCGGGGCGCTATGA	3649
QY	1578	CCAGTGCACAGTGTGAGTGTGATAGTTGTGGAAGAGTCAAAAGGCTCTCTCAAGCGT	1637	QY	2658	CATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACTATCTCTGGGCATCGGCAC	2717
Db	2570	CCAGTGCACAGTGTGAGTGTGATAGTTGTGGAAGAGTCAAAAGGCTCTCTCAAGCGT	2629	Db	3650	CATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACTATCTCTGGGCATCGGCAC	3709
QY	1638	ATTCAACAAAGGGCTGAAGGATGCCAGAAAGTACCCCATTTGTATGGGATCTGATCTGGG	1697	QY	2718	AGTCTTGACCAAGCGGAGAGCGGTGGAGCGCACTCGTCTGCTCGCCACCGCTACGCC	2777
Db	2630	ATTCAACAAAGGGCTGAAGGATGCCAGAAAGTACCCCATTTGTATGGGATCTGATCTGGG	2689	Db	3710	AGTCTTGACCAAGCGGAGAGCGGTGGAGCGCACTCGTCTGCTCGCCACCGCTACGCC	3769
QY	1698	GCCTCGGTGCACATGCTTTACATGTTTGTAGTGTGAGGTTAAAAAGCTTAGGCCCCCGG	1757	QY	2778	TCCGGGATCGCTCACCGTGCACATCCAAACATCGAGGAGTGGCTCTCTCCAGCACTGG	2837
Db	2690	GCCTCGGTGCACATGCTTTACATGTTTGTAGTGTGAGGTTAAAAAGCTTAGGCCCCCGG	2749	Db	3770	TCCGGGATCGCTCACCGTGCACATCCAAACATCGAGGAGTGGCTCTCTCCAGCACTGG	3829
QY	1758	AACCAAGGGAGCTGTTTCTCTTGAAGAACACGATATACCATGGCGCTTATACGGC	1817	QY	2838	AGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATCAAGGGGGGAGGACCT	2897
Db	2750	AACCAAGGGAGCTGTTTCTCTTGAAGAACACGATATACCATGGCGCTTATACGGC	2809	Db	3830	AGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATCAAGGGGGGAGGACCT	3889
QY	1818	CTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACTAGCCCTCACAGGCGGGGA	1877	QY	2898	CATTTTTCGCAATCCAAAGAAATGTGATGAGTCCGCGGAAAGCTGTCTCGGCTCGG	2957
Db	2810	CTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACTAGCCCTCACAGGCGGGGA	2869	Db	3890	CATTTTTCGCAATCCAAAGAAATGTGATGAGTCCGCGGAAAGCTGTCTCGGCTCGG	3949
QY	1878	CAGGAACCAAGTGTGAGGGGAGGTCCAGTGTGTCTCCACCGCAACAAATCTTCTGGC	1937	QY	2958	ACTCAATCTGTAGCATATACCGGGGCTTGTATGTCNTCCGTCATACCACTTAGCGGGA	3017
Db	2870	CAGGAACCAAGTGTGAGGGGAGGTCCAGTGTGTCTCCACCGCAACAAATCTTCTGGC	2929	Db	3950	ACTCAATCTGTAGCATATACCGGGGCTTGTATGTCNTCCGTCATACCACTTAGCGGGA	4009
QY	1938	GACCTGCTCAATGAGCGGTGTGACACTGTCTATCATGTGCGCGCTCAAGACCTTGC	1997	QY	3018	CGTCAATCTGTAGCAACCGACGCTCTAATGACCGGCTTTACCGGCGATTTTCGACTCAGT	3077
Db	2930	GACCTGCTCAATGAGCGGTGTGACACTGTCTATCATGTGCGCGCTCAAGACCTTGC	2989	Db	4010	CGTCAATCTGTAGCAACCGACGCTCTAATGACCGGCTTTACCGGCGATTTTCGACTCAGT	4069
QY	1998	CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACCAAGGACCTCGTGGGTG	2057	QY	3078	GATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAGCTTGGACCCGACCTTCAC	3137
Db	2990	CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACCAAGGACCTCGTGGGTG	3049	Db	4070	GATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAGCTTGGACCCGACCTTCAC	4129
QY	2058	GCAAGGCCCCCGGGGCGTTCCTTGACACCATGCACTTCGCGGAGCTCGGACCTTTA	2117	QY	3138	CATTTCAGACGAGCAGCGGTGCCACAGACGCGGTGTACGCTCGACGCGCGGAGGAGAC	3197
Db	3050	GCAAGGCCCCCGGGGCGTTCCTTGACACCATGCACTTCGCGGAGCTCGGACCTTTA	3109	Db	4130	CATTTCAGACGAGCAGCGGTGCCACAGACGCGGTGTACGCTCGACGCGCGGAGGAGAC	4189
QY	2118	CTTGTGTACAGAGGATGCGATGTCTATTCGGGTGCGCGGGGCGACAGAGGGGAG	2177	QY	3198	TGGTAGGGGCGAGGATGGGCAITTTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT	3257
Db	3110	CTTGTGTACAGAGGATGCGATGTCTATTCGGGTGCGCGGGGCGACAGAGGGGAG	3169	Db	4190	TGGTAGGGGCGAGGATGGGCAITTTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT	4249
QY	2178	CCTACTCTCCCCAGGCGCGTCTCTACTTGAAGGCTTTCGGGCGGTGCCACTGCTCTG	2237	QY	3258	GTTTCGATTCCTCGGTTCTGTGGAGTGTATGACGCGGGCTGTGCTTGGTAGAGTTCAC	3317
Db	3170	CCTACTCTCCCCAGGCGCGTCTCTACTTGAAGGCTTTCGGGCGGTGCCACTGCTCTG	3229	Db	4250	GTTTCGATTCCTCGGTTCTGTGGAGTGTATGACGCGGGCTGTGCTTGGTAGAGTTCAC	4309
QY	2238	CCCTCTGGGCGACGCTGTGGGATCTTTCGGGCTCGGCTGTGACCCCGAGGGGTTGGAA	2297	QY	3318	SCCCSCCGAGACCTCAGTTAGTTTGGGCTTACCTTAAACACACAGGGTTGCCGCTGTG	3377
Db	3230	CCCTCTGGGCGACGCTGTGGGATCTTTCGGGCTCGGCTGTGACCCCGAGGGGTTGGAA	3289	Db	4310	SCCCSCCGAGACCTCAGTTAGTTTGGGCTTACCTTAAACACACAGGGTTGCCGCTGTG	4369
QY	2298	GGCGGTGACCTTGTACCGCTCGAGTCTATGGAACCACTATGCGGTCCCGGCTTTCAC	2357	QY	3378	CCAGGACCATCTGGAGTTCTGGGAGAGCTTTTACAGGCTTACCCCACTAGAGCCGCA	3437
Db	3290	GGCGGTGACCTTGTACCGCTCGAGTCTATGGAACCACTATGCGGTCCCGGCTTTCAC	3349	Db	4370	CCAGGACCATCTGGAGTTCTGGGAGAGCTTTTACAGGCTTACCCCACTAGAGCCGCA	4429
QY	2358	GGACAACTGTCCTCCCGCGGTACCGCAGACATTTCCAGGTGGGCCATCTACAGCCGC	2417	QY	3438	TTTCTTGTCCAGACTTAAGCAGGAGGAGACACTTCCCTTACCTGTTAGCATACAGGC	3497
Db	3350	GGACAACTGTCCTCCCGCGGTACCGCAGACATTTCCAGGTGGGCCATCTACAGCCGC	3409	Db	4430	TTTCTTGTCCAGACTTAAGCAGGAGGAGACACTTCCCTTACCTGTTAGCATACAGGC	4489
QY	2418	TACTGTTAGCGCAGAGACCTAAGTCCGCTCGGTATGAGCCCGGATTAAGT	2477	QY	3498	TACGTTGTGCGCAGGCTCAGGCTCCACCTCCATCTGTGGGACCAAAATGTGAAGTGTCT	3557
Db	3410	TACTGTTAGCGCAGAGACCTAAGTCCGCTCGGTATGAGCCCGGATTAAGT	3469	Db	4490	TACGTTGTGCGCAGGCTCAGGCTCCACCTCCATCTGTGGGACCAAAATGTGAAGTGTCT	4549
				QY	3558	CATACGGGTAAAGCCTACGCTGCAAGGCGCCAAAGCCCTTGTGTATAGGCTGGAGCCCT	3617

|||||  
4550 CATAGGCTAAGCCCTACCGTGCACGGCCCAACGCCCTCTGTTATAGGCTGGAGCCGT 4609  
QY  
3618 TCAAAACGAGGTACTACACACACACCCCAATAACCAATACATCATGGCATGATGTCGCG 3677  
Db  
4610 TCAAAACGAGGTACTACACACACACCCCAATAACCAATACATCATGGCATGATGTCGCG 4669  
QY  
3678 TGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGAGTCTTAGCAGCTCTGSC 3737  
Db  
4670 TGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGAGTCTTAGCAGCTCTGSC 4729  
QY  
3738 CGCGTATTCCTGACAAACAGGACAGCTGCTCATTTGTGGCAGGATCATCTTGTCCGAAA 3797  
Db  
4730 CGCGTATTCCTGACAAACAGGACAGCTGCTCATTTGTGGCAGGATCATCTTGTCCGAAA 4789  
QY  
3798 GCGCGCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTGATGAGATGGAAGAGTG 3857  
Db  
4790 GCGCGCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTGATGAGATGGAAGAGTG 4849  
QY  
3858 GCGCTACACCTCCCTTATCATCGAAACAGGAATGACCTGCGCGGAACAAATTCAAACAGAA 3917  
Db  
4850 GCGCTACACCTCCCTTATCATCGAAACAGGAATGACCTGCGCGGAACAAATTCAAACAGAA 4909  
QY  
3918 GGCAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTGCTCCCTGGTGA 3977  
Db  
4910 GGCAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTGCTCCCTGGTGA 4969  
QY  
3978 ATCCAAAGTGGCGAACCTCGAAGCCCTCTGGGCAAGCATATGTGGAATTTATCAGCGG 4037  
Db  
4970 ATCCAAAGTGGCGAACCTCGAAGCCCTCTGGGCAAGCATATGTGGAATTTATCAGCGG 5029  
QY  
4038 GATCAATATTTAGCAGGCTTGTCACTGCTGCTGGCAACCGCGGATGACATCTGAT 4097  
Db  
5030 GATCAATATTTAGCAGGCTTGTCACTGCTGCTGGCAACCGCGGATGACATCTGAT 5089  
QY  
4098 GGCATTACAGCCTCTATCACAGCCGCTCACCAACCAACATACCCCTCTGTTTAAACAT 4157  
Db  
5090 GGCATTACAGCCTCTATCACAGCCGCTCACCAACCAACATACCCCTCTGTTTAAACAT 5149  
QY  
4158 CTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGCGTGTCTGCTTTCGTAGG 4217  
Db  
5150 CTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGCGTGTCTGCTTTCGTAGG 5209  
QY  
4218 CCGCGCATCGTGGAGCGGCTGTTGGAGCATAGCGCTTGGAGGTCCTTGTGATAT 4277  
Db  
5210 CCGCGCATCGTGGAGCGGCTGTTGGAGCATAGCGCTTGGAGGTCCTTGTGATAT 5269  
QY  
4278 TTTGGCAGGTTATGGAGCAGGGGTGGCAGCGGCTCGTGGCTTTAAGGTATGAGCGG 4337  
Db  
5270 TTTGGCAGGTTATGGAGCAGGGGTGGCAGCGGCTCGTGGCTTTAAGGTATGAGCGG 5329  
QY  
4338 CGAGATGCCCTCAACGAGGACCTGGTTAACTTCTTCTCTGCTATCTCTCCCTGGCG 4397  
Db  
5330 CGAGATGCCCTCAACGAGGACCTGGTTAACTTCTTCTCTGCTATCTCTCCCTGGCG 5389  
QY  
4398 CCTAGTCTCGGGTCTGTCGACGAGTACTGCTCGGACGTCGCTGGCGACGTCGGGCCAGGGAGGG 4457  
Db  
5390 CCTAGTCTCGGGTCTGTCGACGAGTACTGCTCGGACGTCGCTGGCGACGTCGGGCCAGGGAGGG 5449  
QY  
4458 GGCCTGTCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGTTAACACGCTCTCC 4517  
Db  
5450 GGCCTGTCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGTTAACACGCTCTCC 5509  
QY  
4518 CACGACTATGTGCTGAGAGGACGCTGACGACGTCGTCACCTAGATCCTCTTAGTCT 4577  
Db  
5510 CACGACTATGTGCTGAGAGGACGCTGACGACGTCGTCACCTAGATCCTCTTAGTCT 5569  
QY  
4578 TACCATCATCAGTCTGAAAGGGTTTCAACAGTGGATCAACAGGACTGCTCCAGCGC 4637  
Db  
5570 TACCATCATCAGTCTGAAAGGGTTTCAACAGTGGATCAACAGGACTGCTCCAGCGC 5629  
QY  
4638 ATGCTCCGCTCGTGGCTAAGAGATCTTTGGGATTGGATATGACCGCTGTTGACTGATTT 4697  
|||||

Db  
5630 ATGCTCCGCTCGTGGCTAAGAGATGTTTGGGATTGGATATGACCGGTGTTGACTGATTT 5689  
QY  
4698 CAAGACCTGGCTCCAGTCCAAGCTCTCTCGCGGATTTGCCGGGAGTCCCTTCTTCTCATG 4757  
Db  
5690 CAAGACCTGGCTCCAGTCCAAGCTCTCTCGCGGATTTGCCGGGAGTCCCTTCTTCTCATG 5749  
QY  
4758 TCAAGCTGGGTACAAGGAGTCTGCGGGGCGACGGCATATGCAAAACACCTGCCCATG 4817  
Db  
5750 TCAAGCTGGGTACAAGGAGTCTGCGGGGCGACGGCATATGCAAAACACCTGCCCATG 5809  
QY  
4818 TCGAGCACAGATCACCGGACATGTGAAAACCGGTTCCATGAGATCGTGGGGCTTAGGAC 4877  
Db  
5810 TCGAGCACAGATCACCGGACATGTGAAAACCGGTTCCATGAGATCGTGGGGCTTAGGAC 5869  
QY  
4878 CTGTAGTAACAACGTGGCATGGAACATTTCCCATTAACGCGTACACACCGGGCCCTGCAC 4937  
Db  
5870 CTGTAGTAACAACGTGGCATGGAACATTTCCCATTAACGCGTACACACCGGGCCCTGCAC 5929  
QY  
4938 GCGCTCCCGCGCGCCAAATTTCTAGGCGCTGTGGGGGTGGTGTGCTGAGGAGTACGT 4997  
Db  
5930 GCGCTCCCGCGCGCCAAATTTCTAGGCGCTGTGGGGGTGGTGTGCTGAGGAGTACGT 5989  
QY  
4998 GGAGGTTACCGGGTGGGGATTTCCACTACGCTGACCGGCATGACCACTGACAACTAAA 5057  
Db  
5990 GGAGGTTACCGGGTGGGGATTTCCACTACGCTGACCGGCATGACCACTGACAACTAAA 6049  
QY  
5058 GTGCCGCTGTCAAGTTCCGGCCCCCGAAATTTCTTCAACAAGTGGATGGGGTGGCTGCA 5117  
Db  
6050 GTGCCGCTGTCAAGTTCCGGCCCCCGAAATTTCTTCAACAAGTGGATGGGGTGGCTGCA 6109  
QY  
5118 CAGGTACGCTCAGCGTGCACACCCCTCTACGGGAGGAGTCACTTCTGTCGGGCT 5177  
Db  
6110 CAGGTACGCTCAGCGTGCACACCCCTCTACGGGAGGAGTCACTTCTGTCGGGCT 6169  
QY  
5178 CAATCAATACCTGTTGGGTACAGCTCCCATGCCAGCCGAAACCGAGTACGAGTGT 5237  
Db  
6170 CAATCAATACCTGTTGGGTACAGCTCCCATGCCAGCCGAAACCGAGTACGAGTGT 6229  
QY  
5238 CACTTCCATGCTACCGACCCCTCCCAATTAACGGCGGAGACGGCTAAGCGTAGGCTGC 5297  
Db  
6230 CACTTCCATGCTACCGACCCCTCCCAATTAACGGCGGAGACGGCTAAGCGTAGGCTGC 6289  
QY  
5298 CAGGGATCTCCCGCTCTCTGGCAGCTCATCAGTACGAGCTGTCTGCGCTTCTCT 5357  
Db  
6290 CAGGGATCTCCCGCTCTCTGGCAGCTCATCAGTATCCAGCTGTCTGCGCTTCTCT 6349  
QY  
5358 GAAGGCAACATCATCTACCCGTCATGACTCCCGGACGCTGACCTCATCAGGCGCAACT 5417  
Db  
6350 GAAGGCAACATCATCTACCCGTCATGACTCCCGGACGCTGACCTCATCAGGCGCAACT 6409  
QY  
5418 CTTGTGGCGCAGGAGATGGCGGGAACATACCCCGTGGAGTCAGAAAAATAAGGTAGT 5477  
Db  
6410 CTTGTGGCGCAGGAGATGGCGGGAACATACCCCGTGGAGTCAGAAAAATAAGGTAGT 6469  
QY  
5478 AATTTTGGATCTTTTCGAGCGCTCCNAGCGAGGAGTGAAGGAGTATCCGTTCC 5537  
Db  
6470 AATTTTGGATCTTTTCGAGCGCTCCNAGCGAGGAGTGAAGGAGTATCCGTTCC 6529  
QY  
5538 GCGGAGATCTCTCGGAGGTCCAGGAAATTTCCCTCGAGCATGCCATATGGCGACGCC 5597  
Db  
6530 GCGGAGATCTCTCGGAGGTCCAGGAAATTTCCCTCGAGCATGCCATATGGCGACGCC 6589  
QY  
5598 GGATTAACAACCTTCACTGTTAGAGTCTTGAAGAACCCCGGACTACGTCCTTCCAGTGT 5657  
Db  
6590 GGATTAACAACCTTCACTGTTAGAGTCTTGAAGAACCCCGGACTACGTCCTTCCAGTGT 6649  
QY  
5658 ACACGGTGTCCATTTGCGGCTGCAAGGCCCTCCGATACCACTCCACGAGGAGAG 5717  
Db  
6650 ACACGGTGTCCATTTGCGGCTGCAAGGCCCTCCGATACCACTCCACGAGGAGAG 6709  
QY  
5718 GACGGTGTCTGTCAGAAATCTACCGTGTCTTCTGCTGGCGGAGTCTGCCACAAAGAC 5777  
Db  
6710 GACGGTGTCTGTCAGAAATCTACCGTGTCTTCTGCTGGCGGAGTCTGCCACAAAGAC 6769  
|||||

QY 5778 CTTGGGAGCTCCGATCGTGGCGTGGACAGCGGACGCGACCGGCTCTCTGACCA 5837  
Db CTTGGGAGCTCCGATCGTGGCGTGGACAGCGGACGCGACCGGCTCTCTGACCA 6829  
QY 5838 GCCCTCCGACGACGCGGAGCGGATCCGATCGTGGAGTGGTACTCCTCATGCCCCCT 5897  
Db GCCCTCCGACGACGCGGAGCGGATCCGATCGTGGAGTGGTACTCCTCATGCCCCCT 6889  
QY 5898 TGAGGGGAGCGGGGGATCCGATCTCAGCGAGGGTCTTGGTCTACCGTAAGCGAGA 5957  
Db TGAGGGGAGCGGGGGATCCGATCTCAGCGAGGGTCTTGGTCTACCGTAAGCGAGA 6949  
QY 5958 GGCTAGTCAGGAGCTGCTGCTGCTGATCTCTACACATGGAAGGCGGCTGATCAC 6017  
Db GGCTAGTCAGGAGCTGCTGCTGCTGATCTCTACACATGGAAGGCGGCTGATCAC 7009  
QY 6018 GCCATGCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCG 6077  
Db GCCATGCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCG 7069  
QY 6078 TCACCAAACTTGTCTATGCTAACAATCTCGAGCGCAAGCTGCGGCGAAGAAGGT 6137  
Db TCACCAAACTTGTCTATGCTAACAATCTCGAGCGCAAGCTGCGGCGAAGAAGGT 7129  
QY 6138 CACCTTTGACAGACTGAGGCTCTGGAGCAACCTACCGGACGCTCAAGGAGATGAA 6197  
Db CACCTTTGACAGACTGAGGCTCTGGAGCAACCTACCGGACGCTCAAGGAGATGAA 7189  
QY 6198 GCGAAGCGCTCCAGTTAAGGTTAACTTCTATCCGTGGAGGAAGCTGTAAAGCTGAC 6257  
Db GCGAAGCGCTCCAGTTAAGGTTAACTTCTATCCGTGGAGGAAGCTGTAAAGCTGAC 7249  
QY 6258 GCCCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAAAGGACGCTCGGAACCTATC 6317  
Db GCCCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAAAGGACGCTCGGAACCTATC 7309  
QY 6318 CAGCAAGCGCTTAACACATCCGCTCGTGGAGGACTTGTGGAGGACACTGAGAC 6377  
Db CAGCAAGCGCTTAACACATCCGCTCGTGGAGGACTTGTGGAGGACACTGAGAC 7369  
QY 6378 ACCAATTGACACCACTCATGGCAAAATGAGTTTCTCGCTCCAAACAGAGAGGG 6437  
Db ACCAATTGACACCACTCATGGCAAAATGAGTTTCTCGCTCCAAACAGAGAGGG 7429  
QY 6438 GGGCGCAAGCGCTCGCTTATCGTATCCGATTTGGGGTTGGGGTTGGTGGAGAA 6497  
Db GGGCGCAAGCGCTCGCTTATCGTATCCGATTTGGGGTTGGGGTTGGTGGAGAA 7489  
QY 6498 AATGGCCCTTTACGATGCTGCTCACCTCCCTCAGGCGCTGATGGGCTCTTCATCGG 6557  
Db AATGGCCCTTTACGATGCTGCTCACCTCCCTCAGGCGCTGATGGGCTCTTCATCGG 7549  
QY 6558 ATTCCAATCTCTCTGACAGCGGTCGAGTTCTGGTGAATGCTCGAAGCGAAGAA 6617  
Db ATTCCAATCTCTCTGACAGCGGTCGAGTTCTGGTGAATGCTCGAAGCGAAGAA 7609  
QY 6618 ATGCGCTATGGCTTCGATATGACACCGCTGTTTGAATCAACGCTCACTGAGATGA 6677  
Db ATGCGCTATGGCTTCGATATGACACCGCTGTTTGAATCAACGCTCACTGAGATGA 7669  
QY 6678 CATCGCTTGGAGGTCAATCTACCAATGTTGACTTGGCCCGGCAAGCGACAGGC 6737  
Db CATCGCTTGGAGGTCAATCTACCAATGTTGACTTGGCCCGGCAAGCGACAGGC 7729  
QY 6738 CATAAGTGGCTTCAAGAGCGCTTTACATCGGGGCCCCCTGACTAATTTAAAGGCA 6797  
Db CATAAGTGGCTTCAAGAGCGCTTTACATCGGGGCCCCCTGACTAATTTAAAGGCA 7789  
QY 6798 GAACTGCGGCTATCGCGGTCCGCGAGGCGGTGCTACTGACGACCGCTCGGTAAATAC 6857  
Db GAACTGCGGCTATCGCGGTCCGCGAGGCGGTGCTACTGACGACCGCTCGGTAAATAC 7849

QY 6858 CCTCACATGTTACTTGAAGCGCGCTGCGGCTGTGAGCTGCGAAGCTCCAGGACTGCAC 6917  
Db CCTCACATGTTACTTGAAGCGCGCTGCGGCTGTGAGCTGCGAAGCTCCAGGACTGCAC 7909  
QY 6918 GATGCTGTATGCGGAGACGACCTTTGCTGTTATCTGTGAAAGCGCGGAGCCCAAGAGA 6977  
Db GATGCTGTATGCGGAGACGACCTTTGCTGTTATCTGTGAAAGCGCGGAGCCCAAGAGA 7969  
QY 6978 CGAGCGAGCTACGCGGCTTACGAGGCTATGACTAGATACTCTGCCCCCCTCGGGGA 7037  
Db CGAGCGAGCTACGCGGCTTACGAGGCTATGACTAGATACTCTGCCCCCCTCGGGGA 8029  
QY 7038 CCGCGCCAAACCAAGATGAGCTTGGAGTTGATTAACATCATGCTCCTCCATGTTGCTAGT 7097  
Db CCGCGCCAAACCAAGATGAGCTTGGAGTTGATTAACATCATGCTCCTCCATGTTGCTAGT 8089  
QY 7098 CGCGCAGATGCTTGGCAAAAGGTTACTATCTCAACCGTGACCCACACCCCT 7157  
Db CGCGCAGATGCTTGGCAAAAGGTTACTATCTCAACCGTGACCCACACCCCT 8149  
QY 7158 TGCGGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACAT 7217  
Db TGCGGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACAT 8209  
QY 7218 CATCATGATGCGGCCACCTTGTGGCAAGGATGATCCTGATGACTCATTTCTTCTCCAT 7277  
Db CATCATGATGCGGCCACCTTGTGGCAAGGATGATCCTGATGACTCATTTCTTCTCCAT 8269  
QY 7278 CTTTCTAGCTCAGGAACAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGGCGCTGTTA 7337  
Db CTTTCTAGCTCAGGAACAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGGCGCTGTTA 8329  
QY 7338 CTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTCCATGCGCTTAGGCAAT 7397  
Db CTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTCCATGCGCTTAGGCAAT 8389  
QY 7398 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCACTCCCTCAGAAACT 7457  
Db TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCACTCCCTCAGAAACT 8449  
QY 7458 TGGGTTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAGTTCGCGCTAGGCTACT 7517  
Db TGGGTTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAGTTCGCGCTAGGCTACT 8509  
QY 7518 GTCCAGGGGGAGGGTGCACCTTGTGGCAAGTACTCTTCACTGGGCGAGTAAGAC 7577  
Db GTCCAGGGGGAGGGTGCACCTTGTGGCAAGTACTCTTCACTGGGCGAGTAAGAC 8569  
QY 7578 CAAGCTCAAACTCACTCCAATCCCGGCTGCGTCCAGTTGGATTATCCAGCTGGTTCT 7637  
Db CAAGCTCAAACTCACTCCAATCCCGGCTGCGTCCAGTTGGATTATCCAGCTGGTTCT 8629  
QY 7638 TGCTGGTTACAGCGGGGAGACATATATCAAGCCTGCTCTGCGCCGACCCCGCTGTT 7697  
Db TGCTGGTTACAGCGGGGAGACATATATCAAGCCTGCTCTGCGCCGACCCCGCTGTT 8689  
QY 7698 CATGCTGTGCTACTCTCTACTTCTGAGGGTAGGATCTATCTACTCCCAACCGATG 7757  
Db CATGCTGTGCTACTCTCTACTTCTGAGGGTAGGATCTATCTACTCCCAACCGATG 8749  
QY 7758 AACGGGACCTTAAACACTCCAGGCAATAGGCCATCTCTGTTTTTTTTTTTTTTTTT 7817  
Db AACGGGACCTTAAACACTCCAGGCAATAGGCCATCTCTGTTTTTTTTTTTTTTTTT 8809  
QY 7818 CTTT 7877  
Db CTTT 8869  
QY 7878 CTTTTTCTTCTTGGTGGCTCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGGCTCG 7937  
Db CTTTTTCTTCTTGGTGGCTCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGGCTCG 8929  
QY 7938 TGAGCCGCTTACTGCTGAGAGAGTCTGATCTGCGCTCTCTGACATCAAGT 7989

```
Db      8930 TGAGCGCTTGACTGACAGAGAGTGCTGATGACTGAGCTCTCTGCGCTCTCTGAGATCAAGT 8981
|||||
RESULT 15
AX937621
LOCUS      AX937621                12315 bp      DNA      linear      PAT 06-JAN-2004
DEFINITION Sequence 1 from Patent WO03091439.
ACCESSION AX937621
VERSION    AX937621.1 GI:40713672
KEYWORDS
SOURCE     unidentified
ORGANISM   unidentified
            unclassified.
REFERENCE  1
AUTHORS    Duggal,R.A., Patrick,A.K., Zhang,J.A. and Zhao,W.A.
TITLE      Reporter-selectable hepatitis c virus replicon
JOURNAL    Patent: WO 03091439-A 1 06-NOV-2003;
            PFIZER INC. (US)
FEATURES   Location/Qualifiers
            source
            1..12315
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
            /note="A cell line wherein the nucleic acid molecule is a
            self replicati ng RNA molecule."
ORIGIN
Query Match      95.2%; Score 7607.2; DB 6; Length 12315;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7609; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 378 GGGCGCGCATGATTGAAACAAGATGATTGACGACGAGTTTCGGCGCGCTTGGGTGGAG 437
Db 1540 GGGCGCGCATGATTGAAACAAGATGATTGACGACGAGTTTCGGCGCGCTTGGGTGGAG 1599
QY 438 AGCGTATTCGGCTATGACTGGGACACAGACAATCGGCTGCTCGATCGCCCGGTTC 497
Db 1600 AGCGTATTCGGCTATGACTGGGACACAGACAATCGGCTGCTCGATCGCCCGGTTC 1659
QY 498 CGCGTGTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACGACCTGTCCGCTGCCCTG 557
Db 1660 CGCGTGTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACGACCTGTCCGCTGCCCTG 1719
QY 558 AATGAATGACGAGACGAGCGAGCGGCTATCGTGGCTGCCACGACGAGCGGCTTCCTGC 617
Db 1720 AATGAATGACGAGACGAGCGAGCGGCTATCGTGGCTGCCACGACGAGCGGCTTCCTGC 1779
QY 618 GCAGCTGTCTGACGCTGTCTACTGAAGCGGAAGGACTGGCTGTATTGGCGGAAGTG 677
Db 1780 GCAGCTGTCTGACGCTGTCTACTGAAGCGGAAGGACTGGCTGTATTGGCGGAAGTG 1839
QY 678 CGGGGACGAGATCTCCTGTCATCTCACTTGTCTTCCGCGAAGATATCCATCATGGCT 737
Db 1840 CGGGGACGAGATCTCCTGTCATCTCACTTGTCTTCCGCGAAGATATCCATCATGGCT 1899
QY 738 GATCAATGCGGGGTGTCATACGTTGATCGGCTACTCGTGGCTGCCACCAACCAAGCG 797
Db 1900 GATCAATGCGGGGTGTCATACGTTGATCGGCTACTCGTGGCTGCCACCAACCAAGCG 1959
QY 798 AAACATCGCATCGAGCGACGATCTACTCGATGGAGCGGCTTGTGCGATCAGGATGAT 857
Db 1960 AAACATCGCATCGAGCGACGATCTACTCGATGGAGCGGCTTGTGCGATCAGGATGAT 2019
QY 858 CTGGACGAAGAGATCAGGGGTTCGGCGCAGCGCAACTGTTCCGCAAGGCTCAAGGCGCG 917
Db 2020 CTGGACGAAGAGATCAGGGGTTCGGCGCAGCGCAACTGTTCCGCAAGGCTCAAGGCGCG 2079
QY 918 ATGCCGACGCGGAGATCTGCTGATGACCAATGACCAATGACCAATGACCAATGACCA 977
Db 2080 ATGCCGACGCGGAGATCTGCTGATGACCAATGACCAATGACCAATGACCAATGACCA 2139
QY 978 GTGGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCGC 1037
|||||
Db      8930 TGAGCGCTTGACTGACAGAGAGTGCTGATGACTGAGCTCTCTGCGCTCTCTGAGATCAAGT 8981
|||||
12140 GTGGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCGC 2199
1038 TATCAGGACATAGCGTTGGCTACCGTGATATTCCTGAAGAGCTTGGCGGCGCAATGGCT 1097
2200 TATCAGGACATAGCGTTGGCTACCGTGATATTCCTGAAGAGCTTGGCGGCGCAATGGCT 2259
1098 GACCGCTTCTCGTGTTCATCGGTTACCGGTTACCGGTTACCGGTTACCGGTTACCGGTTAC 1157
2260 GACCGCTTCTCGTGTTCATCGGTTACCGGTTACCGGTTACCGGTTACCGGTTACCGGTTAC 2319
1158 CGCGTCTTGTGAGAGTTCTTCTGAGTTTAAACAGACCAACAGGTTTCCCTCTAGCGGA 1217
2320 CGCGTCTTGTGAGAGTTCTTCTGAGTTTAAACAGACCAACAGGTTTCCCTCTAGCGGA 2379
1218 TCAATTCGCGCGCTCTCCCTCCCGCGCGCTTACTGCGCGGAGCGCTTGGGAATA 1277
2380 TCAATTCGCGCGCTCTCCCTCCCGCGCGCTTACTGCGCGGAGCGCTTGGGAATA 2439
1278 AGGCGGTGTGGTGTTCATATGTTATTTTCCACCATATTTGCGCTCTTTTGGCAATGT 1337
2440 AGGCGGTGTGGTGTTCATATGTTATTTTCCACCATATTTGCGCTCTTTTGGCAATGT 2499
1338 GAGGCGCGGAAACCTGGCGCTCTCTTGTGAGGACATTTCTAGGCTTCTTCCCTCT 1397
2500 GAGGCGCGGAAACCTGGCGCTCTCTTGTGAGGACATTTCTAGGCTTCTTCCCTCT 2559
1398 CGCCAAAGGAATGCAAGTCTCTTGAATTCGTGAAGAGAGTTCCTCTGGAAGCTTC 1457
2560 CGCCAAAGGAATGCAAGTCTCTTGAATTCGTGAAGAGAGTTCCTCTGGAAGCTTC 2619
1458 TTGAAGACAAACAAACGCTGTGTAGGACCTTTTGTGAGGACGAGAACCCCGCTTGGGGA 1517
2620 TTGAAGACAAACAAACGCTGTGTAGGACCTTTTGTGAGGACGAGAACCCCGCTTGGGGA 2679
1518 CAGGTGCTCTCGGCGCAAAAGCGCTGTATTAAGATACCTGTGAAGGCGGCAACCC 1577
2680 CAGGTGCTCTCGGCGCAAAAGCGCTGTATTAAGATACCTGTGAAGGCGGCAACCC 2739
1578 CCAGTGCCACAGTGTGAGTTGATAGTTGTGAAAGAGTCAATAGCTCTCTCAAGCGT 1637
2740 CCAGTGCCACAGTGTGAGTTGATAGTTGTGAAAGAGTCAATAGCTCTCTCAAGCGT 2799
1638 ATTCAACAAAGGGCTGAAGATGCCAGAAAGTACCCCATTTGTATGGATCTGATCGG 1697
2800 ATTCAACAAAGGGCTGAAGATGCCAGAAAGTACCCCATTTGTATGGATCTGATCGG 2859
1698 GCCTCGGTGCATGCTTTACATGTTTGTAGTGTAAAGAGCTTACGAGGCTTACGAGG 1757
2860 GCCTCGGTGCATGCTTTACATGTTTGTAGTGTAAAGAGCTTACGAGGCTTACGAGG 2919
1758 AACCCAGGGGAGCTGTTTCTTCTTGAAGAAACAGATATACCATGGCGCTTATACGCG 1817
2920 AACCCAGGGGAGCTGTTTCTTCTTGAAGAAACAGATATACCATGGCGCTTATACGCG 2979
1818 CTACTCCCAACAGACGAGGCTTCTTGGCTGTGATCATCTAGCTTCAAGGCGGGA 1877
2980 CTACTCCCAACAGACGAGGCTTCTTGGCTGTGATCATCTAGCTTCAAGGCGGGA 3039
1878 CAGGAACAGAGTTCGAGGGGAGGCTTCCAGGCTTCCACCGCAACACATCTTCTTGGC 1937
3040 CAGGAACAGAGTTCGAGGGGAGGCTTCCAGGCTTCCACCGCAACACATCTTCTTGGC 3099
1938 GACCTGCGCTCAATGGCGTGTGTGTGACTGTCTATCATGTGCGGCTCAAGACCTTGC 1997
3100 GACCTGCGCTCAATGGCGTGTGTGTGACTGTCTATCATGTGCGGCTCAAGACCTTGC 3159
1998 CGGCGCAAGGGGCGCAATCACCAGATGTACCAATGTGACCAAGGACCTCGTGGGTG 2057
3160 CGGCGCAAGGGGCGCAATCACCAGATGTACCAATGTGACCAAGGACCTCGTGGGTG 3219
2058 GCAAGCGCGCGCGCGGCTTCTTGTGACCATGTGACCTGCGGAGCTTGGACCTTTA 2117
|||||
```



QY 4278 TTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGCCCTTTAAGTTCATGAGGG 4337  
Db 5440 TTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGCCCTTTAAGGTCATGAGCGG 5499  
QY 4338 CGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCTCTACTATCCTCTCCCTCGCGC 4397  
Db 5500 CGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCTCTACTATCCTCTCCCTCGCGC 5559  
QY 4398 CCTAGTCGTGGGGTCTGTGCGCAGCGATACCTGCTGCGACGCTGGGCGCCAGGGAGGG 4457  
Db 5560 CCTAGTCGTGGGGTCTGTGCGCAGCGATACCTGCTGCGACGCTGGGCGCCAGGGAGGG 5619  
QY 4458 GGGTGTGCACTGAGTAACCGGCTGATAGCTTCGCTTCGGGGTAAACACGCTCCGCC 4517  
Db 5620 GGGTGTGCACTGAGTAACCGGCTGATAGCTTCGCTTCGGGGTAAACACGCTCCGCC 5679  
QY 4518 CAGCACTATGTGCTGAGAGCGACGCTGAGCAGCGTGTCACTCAGATCCCTCTAGTCT 4577  
Db 5680 CAGCACTATGTGCTGAGAGCGACGCTGAGCAGCGTGTCACTCAGATCCCTCTAGTCT 5739  
QY 4578 TACCATCAGCTGCTGAGAGGCTTACAGTGGATCAACGAGAGCTGCTCCAGCC 4637  
Db 5740 TACCATCAGCTGCTGAGAGGCTTACAGTGGATCAACGAGAGCTGCTCCAGCC 5799  
QY 4638 ATGCTCCGGCTCGTGGCTAAGAGATGTTGGGATTGGATATGACGCTGTGACTGATTT 4697  
Db 5800 ATGCTCCGGCTCGTGGCTAAGAGATGTTGGGATTGGATATGACGCTGTGACTGATTT 5859  
QY 4698 CAAGACTGCTCAAGTCCAGTCTCTGCGCGGATTCGCGGAGTCCCTTCTCTCATG 4757  
Db 5860 CAAGACTGCTCAAGTCCAGTCTCTGCGCGGATTCGCGGAGTCCCTTCTCTCATG 5919  
QY 4758 TCACGCTGGCTCAAGGAGCTTGGCGGGGAGCGGATCATGCAACACCTGCGCATG 4817  
Db 5920 TCACGCTGGGTAAGAGGAGCTTGGCGGGGAGCGGATCATGCAACACCTGCGCATG 5979  
QY 4818 TGGAGCACAGATCACCGGACATGTGAAAAACGGTTCATGAGGATCGTGGGCGCTAGAC 4877  
Db 5980 TGGAGCACAGATCACCGGACATGTGAAAAACGGTTCATGAGGATCGTGGGCGCTAGAC 6039  
QY 4878 CTGTAGTAACAGTGGATGGAACATTTCCCATTTAACCGCTACACACGCGCCCTGCAC 4937  
Db 6040 CTGTAGTAACAGTGGATGGAACATTTCCCATTTAACCGCTACACACGCGCCCTGCAC 6099  
QY 4938 GGCCTCCCGGCGCAAAATTTCTAGGCGCTGTGCGGGGTGCTGCTGAGGAGTACGT 4997  
Db 6100 GGCCTCCCGGCGCAAAATTTCTAGGCGCTGTGCGGGGTGCTGCTGAGGAGTACGT 6159  
QY 4998 GGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGATGACACCTGACAACTGAA 5057  
Db 6160 GGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGATGACACCTGACAACTGAA 6219  
QY 5058 GTGCCCGTGCAGTTCCGGCCCGGAAATTTTTCACAGAGTGGATGGGGTGGGTTGCA 5117  
Db 6220 GTGCCCGTGCAGTTCCGGCCCGGAAATTTTTCACAGAGTGGATGGGGTGGGTTGCA 6279  
QY 5118 CAGGTACGCTCCAGCGTGCAAACCCCTCTACCGGAGGAGTCAATTCCTGCTGGGCT 5177  
Db 6280 CAGGTACGCTCCAGCGTGCAAACCCCTCTACCGGAGGAGTCAATTCCTGCTGGGCT 6339  
QY 5178 CAATCAATACCTGGTGGGTTCAGCTCCCATCGAGCGCGGAAACCGGATGACGTAGCTGCT 5237  
Db 6340 CAATCAATACCTGGTGGGTTCAGCTCCCATCGAGCGCGGAAACCGGATGACGTAGCTGCT 6399  
QY 5238 CACTTCATGCTCACCGAACCCTCCCAATTCAGGCGGAGACGGCTAAGCGTAGGCTGGC 5297  
Db 6400 CACTTCATGCTCACCGAACCCTCCCAATTCAGGCGGAGACGGCTAAGCGTAGGCTGGC 6459  
QY 5298 CAGGGATCTCCCCCTCTCTGGCAGCTCATCAGTACGAGCTGTGCGCTTCTT 5357  
Db 6460 CAGGGATCTCCCCCTCTCTGGCAGCTCATCAGTACGAGCTGTGCGCTTCTT 6519  
QY 5358 GAAGGCAACATGCACTACCGCTCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT 5417

Db 6520 GAAGGCAACATGCACTACCGCTCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT 6579  
QY 5418 CCTGTGCGCGCAGGAGATGGCGGGAACATCACCCGCTGGAGTCAGAAAATAGGTAGT 5477  
Db 6580 CCTGTGCGCGCAGGAGATGGCGGGAACATCACCCGCTGGAGTCAGAAAATAGGTAGT 6639  
QY 5478 AATTTTGGACTCTTTTCAGCGCTTCCAAAGCGGAGGAGATGAGAGGAAATATCCGTTCC 5537  
Db 6640 AATTTTGGACTCTTTTCAGCGCTTCCAAAGCGGAGGAGATGAGAGGAAATATCCGTTCC 6699  
QY 5538 GCGGAGATCCTGCGGAGTCCAGGAAATTCCTCGAGCGATGCCCATATGCGGACGCGC 5597  
Db 6700 GCGGAGATCCTGCGGAGTCCAGGAAATTCCTCGAGCGATGCCCATATGCGGACGCGC 6759  
QY 5598 GGATTACAAACCTTCCACTGTAGAGTCTCGMAGGACCCGAGCTACGCTCCCTCAGTGGT 5657  
Db 6760 GGATTACAAACCTTCCACTGTAGAGTCTCGMAGGACCCGAGCTACGCTCCCTCAGTGGT 6819  
QY 5658 ACACGGGTGCTAATTTGCGGCTGCCAAGGCCCTCGATACCACTCCACGGAGGAAGAG 5717  
Db 6820 ACACGGGTGCTAATTTGCGGCTGCCAAGGCCCTCGATACCACTCCACGGAGGAAGAG 6879  
QY 5718 GACGGTGTCTCTCAGAACTACCGTGTCTTCTGCTTGGCGAGCTCGCCACAAGAC 5777  
Db 6880 GACGGTGTCTCTCAGAACTACCGTGTCTTCTGCTTGGCGAGCTCGCCACAAGAC 6939  
QY 5778 CTTTCGGCAGCTCCGAATTCGTGCGCGCTCGACAGCGCACGCAACGCGCTCTCTGACCA 5837  
Db 6940 CTTTCGGCAGCTCCGAATTCGTGCGCGCTCGACAGCGCACGCAACGCGCTCTCTGACCA 6999  
QY 5838 GCCTTCGAGCAGCGCGACGCGGATTCGACGTTGAGTCTCTCTCTCTATGCGGCGCT 5897  
Db 7000 GCCTTCGAGCAGCGCGGATTCGACGTTGAGTCTCTCTCTCTATGCGGCGCT 7059  
QY 5898 TGAGGGGAGCGCGGGATTCGCGATCTCAGCAGCGGTCTTGGTCTACCGTAAGCAGGA 5957  
Db 7060 TGAGGGGAGCGCGGGATTCGCGATCTCAGCAGCGGTCTTGGTCTACCGTAAGCAGGA 7119  
QY 5958 GGCTAGTGAGGAGCTGCTCTGCTGCTCGATGTCTACATGACAGCGCGCTCTGATCAC 6017  
Db 7120 GGCTAGTGAGGAGCTGCTCTGCTGCTCGATGTCTACATGACAGCGCGCTCTGATCAC 7179  
QY 6018 GCCATCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCGG 6077  
Db 7180 GCCATCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCGG 7239  
QY 6078 TCACCAAACTTGGTCTATGCTACAACTCTCGCAGCGCAAGCTCGCGGAGAGAGGT 6137  
Db 7240 TCACCAAACTTGGTCTATGCTACAACTCTCGCAGCGCAAGCTCGCGGAGAGAGGT 7299  
QY 6138 CACTTTGACAGACTCGAGTCTGAGACCACTACCGGAGCTCTCAAGGAGATGAA 6197  
Db 7300 CACTTTGACAGACTCGAGTCTGAGACCACTACCGGAGCTCTCAAGGAGATGAA 7359  
QY 6198 GCGGAAGCGCTCAGTTAAGCTAAAGCTTCTATCCGTGGAGGAGCTGTAAGCTGAC 6257  
Db 7360 GCGGAAGCGCTCAGTTAAGCTAAAGCTTCTATCCGTGGAGGAGCTGTAAGCTGAC 7419  
QY 6258 GCCCCACATTCGCGAGTCTAATTTGGCTATGGGCAAGGAGCTCGGNAACCTATC 6317  
Db 7420 GCCCCACATTCGCGAGTCTAATTTGGCTATGGGCAAGGAGCTCGGNAACCTATC 7479  
QY 6318 CAGCAAGGCGGTTAAACCATCCGCTCCGTGTGGAGGACTTGTGGAAGACACTGAGAC 6377  
Db 7480 CAGCAAGGCGGTTAAACCATCCGCTCCGTGTGGAGGACTTGTGGAAGACACTGAGAC 7539  
QY 6378 ACCAATGACACCAATCATGCGCAAAAATAGGTTTCTCGTCCCAACCGAGAGGG 6437  
Db 7540 ACCAATGACACCAATCATGCGCAAAAATAGGTTTCTCGTCCCAACCGAGAGGG 7599  
QY 6438 GGGCGCAGCCAGCTCGGCTTATCGTATCCAGATTTGGGGGTTCGTGTGCGAGAA 6497

Db 7600 GGGCGGCAAGCCAGCTCGCTTATCGTATCCAGATTGGGGGTTGCTGTGCGAGAA 7659  
 Qy 6498 AATGGCCCTTTACGATGFGTCTCCACCTCCCTCAGGCCGTGATGGGCTCTTATACGG 6557  
 Db 7660 AATGGCCCTTTACGATGFGTCTCCACCTCCCTCAGGCCGTGATGGGCTCTTATACGG 7719  
 Qy 6558 ATTCCCATACTCTCTCGACAGCGGTCGAGTCTCTGGTGAATGCTCGAAGCGAAGAA 6617  
 Db 7720 ATTCCCATACTCTCTCGACAGCGGTCGAGTCTCTGGTGAATGCTCGAAGCGAAGAA 7779  
 Qy 6618 ATGCCCTATGGCTTCGATATGACACCCGCTGTTTTCAGCTCAACGGTCACTGAGAATGA 6677  
 Db 7780 ATGCCCTATGGCTTCGATATGACACCCGCTGTTTTCAGCTCAACGGTCACTGAGAATGA 7839  
 Qy 6678 CATCGTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCCCCCGGAAAGCGACAGGC 6737  
 Db 7840 CATCGTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCCCCCGGAAAGCGACAGGC 7899  
 Qy 6738 CATAAGGTGCGTCAAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAAGGGCA 6797  
 Db 7900 CATAAGGTGCGTCAAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAAGGGCA 7959  
 Qy 6798 GAACTGCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGACAGCTCGGTAAATAC 6857  
 Db 7960 GAACTGCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGACAGCTCGGTAAATAC 8019  
 Qy 6858 CTTACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGTGGGAAGCTCCAGGACTGCAC 6917  
 Db 8020 CTTACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGTGGGAAGCTCCAGGACTGCAC 8079  
 Qy 6918 GATGCTCTATCGGAGACGACTTGTGTTATCTGTGAAGCGGGGACCCCAAGAGA 6977  
 Db 8080 GATGCTCTATCGGAGACGACTTGTGTTATCTGTGAAGCGGGGACCCCAAGAGA 8139  
 Qy 6978 CGAGCGAGCCTTACGGGCGCTTACGAGGCTATGACTAGATCTGTGCCCCCTGGGGA 7037  
 Db 8140 CGAGCGAGCCTTACGGGCGCTTACGAGGCTATGACTAGATCTGTGCCCCCTGGGGA 8199  
 Qy 7038 CCGGCGCCAAACAGAAATAGGACTTGGAGTTGATACATCATGCTCTCTCAATGTCTAGT 7097  
 Db 8200 CCGGCGCCAAACAGAAATAGGACTTGGAGTTGATACATCATGCTCTCTCAATGTCTAGT 8259  
 Qy 7098 CGCGCAGCATGATCTGGCAAAAGGTGTACTATCTCACCCGTGACCCACCCACCCCT 7157  
 Db 8260 CGCGCAGCATGATCTGGCAAAAGGTGTACTATCTCACCCGTGACCCACCCACCCCT 8319  
 Qy 7158 TCGCGCGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCAACAT 7217  
 Db 8320 TCGCGCGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCAACAT 8379  
 Qy 7218 CATCATGTATGGCCGACCTTGTGGCAAGGATGATCTGTAGTCACTTCTTCTCCAT 7277  
 Db 8380 CATCATGTATGGCCGACCTTGTGGCAAGGATGATCTGTAGTCACTTCTTCTCCAT 8439  
 Qy 7278 CTTCTTAGCTCAGGAACAACCTTGAAGGACCTTAGATCTCAGATCTACGGGCGCTGTTA 7337  
 Db 8440 CTTCTTAGCTCAGGAACAACCTTGAAGGACCTTAGATCTCAGATCTACGGGCGCTGTTA 8499  
 Qy 7338 CTTCCATTGAGCCACTTGACCTACCTCAGATCAATCAAGACTCCATGGCCTTAGGCAAT 7397  
 Db 8500 CTTCCATTGAGCCACTTGACCTACCTCAGATCAATCAAGACTCCATGGCCTTAGGCAAT 8559  
 Qy 7398 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAACCT 7457  
 Db 8560 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAACCT 8619  
 Qy 7458 TGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGTGTCCGGCTAGGCTACT 7517  
 Db 8620 TGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGTGTCCGGCTAGGCTACT 8679  
 Qy 7518 GTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTCACTGGGCGAGTAGGAC 7577  
 Db 8680 GTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTCACTGGGCGAGTAGGAC 8739

Qy 7578 CARGCTCAAACTCACTCCAATCCGGCTGGCTCCAGTTGGATTATCCAGCTGGTTCGT 7637  
 Db 8740 CARGCTCAAACTCACTCCAATCCGGCTGGCTCCAGTTGGATTATCCAGCTGGTTCGT 8799  
 Qy 7638 TGCTGGTTACAGCGGGGAGACATATATACAGGCTGTCTCGTGCACCCCGCTGGTT 7697  
 Db 8800 TGCTGGTTACAGCGGGGAGACATATATACAGGCTGTCTCGTGCACCCCGCTGGTT 8859  
 Qy 7698 CATGTGGTGCCTACTCCTACTCTTCTGTAGGGTGGCATCTATCTACTCCCAACCGATG 7757  
 Db 8860 CATGTGGTGCCTACTCCTACTCTTCTGTAGGGTGGCATCTATCTACTCCCAACCGATG 8919  
 Qy 7758 AACGGGACCTAAACACTCCAGGCGCAATAGGCCATCTCTGTTTTTTCCTTTTTTTTT 7817  
 Db 8920 AACGGGACCTAAACACTCCAGGCGCAATAGGCCATCTCTGTTTTTTCCTTTTTTTTT 8979  
 Qy 7818 CTTT 7877  
 Db 8980 CTTT 9039  
 Qy 7878 CTTTTCTTTCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAGSTCCG 7937  
 Db 9040 CTTTTCTTTCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAGSTCCG 9099  
 Qy 7938 TGAGCCGCTTGACCTGCAGAGAGTGTCTGATCTGGGCTCTCTGCAGATCAAGT 7989  
 Db 9100 TGAGCCGCTTGACCTGCAGAGAGTGTCTGATCTGGGCTCTCTGCAGATCAAGT 9151

Search completed: December 9, 2004, 12:31:02  
 Job time : 32078 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 18:04:05 ; Search time 3311 Seconds

(without alignments)  
12666.151 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccccgcattgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_23Sep04:\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	7989	100.0	7989	6	AAD25322
2	7989	100.0	11313	12	ADP86264
3	7987.4	100.0	7989	3	AAA98968
4	7987.4	100.0	7989	6	AAD25326
5	7987.4	100.0	7989	12	ADJ57845
6	7987.4	100.0	7992	6	AAL47276
7	7987.4	100.0	10690	6	ABK91412
8	7987.4	100.0	10690	9	ACA61697
9	7987.4	100.0	10690	10	ADC83762
10	7987.4	100.0	11313	12	ADP86271
11	7985.8	100.0	7992	6	AAL47281
12	7985.8	100.0	10690	6	ABK91448
13	7985.8	100.0	10690	6	ABK91435
14	7985.8	100.0	10690	6	ABK91243
15	7985.8	100.0	10690	6	ABK91434
16	7985.8	100.0	11313	12	ADP86265
17	7984.2	99.9	7989	6	AAD25325
18	7984.2	99.9	7989	10	ADD93734
19	7984.2	99.9	10690	6	ABK91440
20	7984.2	99.9	11313	12	ADP86272
21	7982.6	99.9	7992	6	AAL47280

22	7982.6	99.9	7992	6	AAL47277
23	7981	99.9	11313	12	ADP86275
24	7979.4	99.9	10690	6	ABK91242
25	7979.4	99.9	11313	12	ADP86273
26	7977.8	99.9	11313	12	ADP86266
27	7976.4	99.8	10691	6	ABK91423
28	7976.2	99.8	11313	12	ADP86268
29	7974.4	99.8	10693	6	ABK91438
30	7973.4	99.8	7987	6	AAD25321
31	7973	99.8	11313	12	ADP86267
32	7972.8	99.8	10693	6	ABK91443
33	7971.8	99.8	7987	6	AAD25329
34	7971.6	99.8	10691	6	ABK91439
35	7970.2	99.8	7987	6	AAD25324
36	7969	99.7	7991	6	AAL47279
37	7965.4	99.7	8001	3	AAA98967
38	7961.6	99.7	7995	6	AAL47278
39	7957	99.6	11313	12	ADP86270
40	7953.8	99.6	11313	12	ADP86269
41	7779.4	97.4	7989	10	ADD93733
42	7713	96.5	11184	12	ADP86274
43	7705	96.4	11184	12	ADP86276
44	7700.2	96.4	11184	12	ADP86277
45	7697	96.3	7848	6	AAD25323

## ALIGNMENTS

## RESULT 1

AAD25322

ID AAD25322 standard; cDNA; 7989 BP.

XX AC AAD25322;

DT 12-MAR-2002 (first entry)

XX DE Hepatitis C virus (HCV) repibBartMan/AvaiI cDNA.

XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;

XX KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver; ss.

XX OS Hepatitis C virus.

XX FH Key

FT CDS Location/Qualifiers

FT 1801..7758

FT /\*tag= a

FT /\*product= "HCVrepibBartMan polyprotein"

FT /\*tag= b

FT /\*note= "Nucleotide creating AvaiI site"

FT WO200189364-A2.

XX 29-NOV-2001.

XX PF 23-MAY-2001; 2001WO-US016822.

XX PR 23-MAY-2000; 2000US-00576989.

XX PA (UNIW ) UNIV WASHINGTON.

XX PI Rice CM, Blight KJ;

XX XX WPI; 2002-066755/09.

XX DR P-PSDB; AAE15717.

XX XX Hepatitis C virus variants having greater transfection efficiency and

PT ability to survive subpassage, useful as a vaccine for immunizing primate

XX PT to the virus, comprise non-naturally occurring viral sequences.

XX PS Claim 44; Page 69-71; 174pp; English.

XX	CC	The invention relates to Hepatitis C virus (HCV) variants which include
CC	CC	polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC	CC	variants that have a transfection efficiency and ability to survive
CC	CC	subpassage greater than HCV that have wild-type polypeptide coding
CC	CC	regions. The polynucleotides of the invention are useful for identifying
CC	CC	a cell line that is permissive for infection with HCV and detecting
CC	CC	replication of HCV in cells of the cell line. They are also useful for
CC	CC	testing a compound for anti-viral properties and for inhibiting HCV
CC	CC	infection. They are also useful for the generation of defined HCV virus
CC	CC	stocks to develop in vitro and in vivo assays for virus neutralisation,
CC	CC	attachment, penetration and entry, structure/function studies on HCV
CC	CC	proteins and RNA elements and identification of new antiviral targets, a
CC	CC	systematic survey of cell culture systems and conditions to identify
CC	CC	those that support wild-type and variant HCV RNA replication and particle
CC	CC	release, production of adaptive HCV variants capable of more efficient
CC	CC	replication in cell culture, production of HCV variants with altered
CC	CC	tissue or species tropism, establishment of alternative animal models for
CC	CC	inhibitor evaluation including those supporting HCV variant replication,
CC	CC	development of cell-free HCV replication assays, production of
CC	CC	immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC	CC	derivatives as possible vaccine candidates, engineering of attenuated or
CC	CC	defective HCV derivatives for expression of heterologous gene products
CC	CC	for gene therapy and vaccine applications and for utilisation of the HCV
CC	CC	glycoproteins for targeted delivery of therapeutic agents to the liver
CC	CC	or other cell types with appropriate receptors. Vaccine comprising these
CC	CC	sequences is useful for inducing immunoprotection to HCV in a primate.
CC	CC	The present sequence is Hepatitis C virus (HCV) replBartMan/Availi cdna
XX	CC	
SQ		Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;
	Query Match	100.0%; Score 7989; DB 6; Length 7989;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 7989; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	GCCAGCCCCGATTGGGGGCGACACTCCACCATGATCACTCCCTGTGAGGAACACTAGT 60
DB	1	GCCAGCCCCGATTGGGGGCGACACTCCACCATGATCACTCCCTGTGAGGAACACTAGT 60
QY	61	TCTTCAACGAGAAACGCTAGCCATGCGCTTAGTATCAGTGTGCTGCAGCCTCCAGGAC 120
DB	61	TCTTCAACGAGAAACGCTAGCCATGCGCTTAGTATCAGTGTGCTGCAGCCTCCAGGAC 120
QY	121	CCCCCTCCCGGGAGAGCCATAGTGGTCTCGCGAACCGGTGAGTACACCGGAATGCCAG 180
DB	121	CCCCCTCCCGGGAGAGCCATAGTGGTCTCGCGAACCGGTGAGTACACCGGAATGCCAG 180
QY	181	GACGACCGGGTCTTTCTTGATCAACCGCTCAATGCTGGAGATTTGGCGGTGCCCCC 240
DB	181	GACGACCGGGTCTTTCTTGATCAACCGCTCAATGCTGGAGATTTGGCGGTGCCCCC 240
QY	241	GGGAGACTGCTAGCCGAGTAGTGTGGGTGCGCGAAGGCTTGTGGTACTGCTGATAGG 300
DB	241	GGGAGACTGCTAGCCGAGTAGTGTGGGTGCGCGAAGGCTTGTGGTACTGCTGATAGG 300
QY	301	GTGCTTGGAGTGCCTCCCGGAGGTCTGTFAGACCGTGACCAATGACGACGATCCTAAAC 360
DB	301	GTGCTTGGAGTGCCTCCCGGAGGTCTGTFAGACCGTGACCAATGACGACGATCCTAAAC 360
QY	361	CTCAAGAAAAACCAAGAGGCGCGCATGATTGAACAAGATGCAATGACGAGGTTCTC 420
DB	361	CTCAAGAAAAACCAAGAGGCGCGCATGATTGAACAAGATGCAATGACGAGGTTCTC 420
QY	421	CGGCGCGTGGGTGGGAGGCTTATCGGCTATGACTGGGCAACAACAGATTCGGCTGCT 480
DB	421	CGGCGCGTGGGTGGGAGGCTTATCGGCTATGACTGGGCAACAACAGATTCGGCTGCT 480
QY	481	CTGATCGCGCGTGTTCGGCTGTACGCGCAGGGCGCCCGTCTTTTGTCAAGACCG 540
DB	481	CTGATCGCGCGTGTTCGGCTGTACGCGCAGGGCGCCCGTCTTTTGTCAAGACCG 540
QY	541	ACCTGTCCGCTGCCCTGAATGAATTCAGGACGAGGACGCGGCTATCGTGGCTGGCA 600

1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCAGAGTTAAAA 1740  
1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCAGAGTTAAAA 1740  
1741 AAGCTCTAGGCCCCCGGAAACCAAGGGGACGTGGTTTTCCTTTGAAAAACACGATAATACC 1800  
1741 AAGCTCTAGGCCCCCGGAAACCAAGGGGACGTGGTTTTCCTTTGAAAAACACGATAATACC 1800  
1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCTTACTTGGCTGCATCATCACT 1860  
1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCTTACTTGGCTGCATCATCACT 1860  
1861 AGCTCACAGGCGGACACAGGACAGTTCGAGGGGAGTCCAAAGTGTCTCACCGGA 1920  
1861 AGCTCACAGGCGGACACAGGACAGTTCGAGGGGAGTCCAAAGTGTCTCACCGGA 1920  
1921 ACACAATCTTCTGGCGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTGC 1980  
1921 ACACAATCTTCTGGCGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTGC 1980  
1981 GGCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
1981 GGCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGCGTTCCTTGCATGTCATCATGTGTGC 2100  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGCGTTCCTTGCATGTCATCATGTGTGC 2100  
2101 GGAGCTCGGACCTTTACTTGTGTCAGAGGATGCGGATGTCATTCGGGTGCGCGCGGG 2160  
2101 GGAGCTCGGACCTTTACTTGTGTCAGAGGATGCGGATGTCATTCGGGTGCGCGCGGG 2160  
2161 GGCGACAGGAGGGAGCTTACTTCTCCCGAGCGCGTCTCTTGAAGGGCTCTTTCG 2220  
2161 GGCGACAGGAGGGAGCTTACTTCTCCCGAGCGCGTCTCTTGAAGGGCTCTTTCG 2220  
2221 GCGGTTCACCTGCTCGCCCTCGGGGACGCTGTGGGACATCTTTCGGGTGCGCGGTGC 2280  
2221 GCGGTTCACCTGCTCGCCCTCGGGGACGCTGTGGGACATCTTTCGGGTGCGCGGTGC 2280  
2281 ACCGAGGGGTTGCGAAGCGGTGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
2281 ACCGAGGGGTTGCGAAGCGGTGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
2341 CGGTCCCCTGTTTACGAGCAACTGCTCCCTCGGGCGGTACCGGAGCACTTCCAGGTG 2400  
2341 CGGTCCCCTGTTTACGAGCAACTGCTCCCTCGGGCGGTACCGGAGCACTTCCAGGTG 2400  
2401 GCGCATCTACAGCGCCCTACTGTAGCGGCAAGACACTAAGGTGCGCGGTGGTATGCA 2460  
2401 GCGCATCTACAGCGCCCTACTGTAGCGGCAAGACACTAAGGTGCGCGGTGGTATGCA 2460  
2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCTCGCGCCACCTAGTTCGGG 2520  
2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCTCGCGCCACCTAGTTCGGG 2520  
2521 GGGTATATGTTAAGGCACATGATCGACCTTAACATCAGAACCGGGGTAAAGACCATC 2580  
2521 GGGTATATGTTAAGGCACATGATCGACCTTAACATCAGAACCGGGGTAAAGACCATC 2580  
2581 ACCAGGGTGCCCGCATCAGTACTCCACTATGCAAGTTCCTTTCGCGACCGGTGGTTC 2640  
2581 ACCAGGGTGCCCGCATCAGTACTCCACTATGCAAGTTCCTTTCGCGACCGGTGGTTC 2640  
2641 TCTGGGGCGCTATGACATATAATATGTATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2641 TCTGGGGCGCTATGACATATAATATGTATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2701 ATCTTGGGATCGGACAGTCTTGAACCAAGCGGAGACGGCTGGAGCGGACTCGTTCGTG 2760  
2701 ATCTTGGGATCGGACAGTCTTGAACCAAGCGGAGACGGCTGGAGCGGACTCGTTCGTG 2760

2761 CTGCCACCCCTAGGCTCCGGGATCGGTCAACCGTGCCACATCOAAACATCGAGGAGTG 2820  
2761 CTGCCACCCCTAGGCTCCGGGATCGGTCAACCGTGCCACATCOAAACATCGAGGAGTG 2820  
2821 GCTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
2821 GCTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAAAGAAATGTGATGAGTCCGCGC 2940  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAAAGAAATGTGATGAGTCCGCGC 2940  
2941 AAGCTGTCCGGCTCGGACTCAATGTCTAGCATATAACGGGGCCCTTGAATATCCGT 3000  
2941 AAGCTGTCCGGCTCGGACTCAATGTCTAGCATATAACGGGGCCCTTGAATATCCGT 3000  
3001 ATACCAACTAGCGGACAGCTATTGTCTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3060  
3001 ATACCAACTAGCGGACAGCTATTGTCTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3060  
3061 GCGATTTTCGACTCAGTCAATACATGATGTGTCAACCCAGACAGTTCGACTTCAGC 3120  
3061 GCGATTTTCGACTCAGTCAATACATGATGTGTCAACCCAGACAGTTCGACTTCAGC 3120  
3121 CTGGACCCGACCTTACCATTTGAGACGACGACCGTGCACAGACGCGGTGTCAACGCTG 3180  
3121 CTGGACCCGACCTTACCATTTGAGACGACGACCGTGCACAGACGCGGTGTCAACGCTG 3180  
3181 CAGCGCGAGGACGAGTGTAGGGGACGATGGGCACTTTACAGGTTTGTGATCTCAGGA 3240  
3181 CAGCGCGAGGACGAGTGTAGGGGACGATGGGCACTTTACAGGTTTGTGATCTCAGGA 3240  
3241 GAAAGCGCCCTCGGGCACTTTCGATTCCTCGGTTCTGTGCGAGTCTATGACGCGGGCTG 3300  
3241 GAAAGCGCCCTCGGGCACTTTCGATTCCTCGGTTCTGTGCGAGTCTATGACGCGGGCTG 3300  
3301 GCTTGTAGAGTACGCGCCCGGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3301 GCTTGTAGAGTACGCGCCCGGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3361 CAGAGGTTCCCGCTGCGCAGGACCATCTGAGAGTCTGGAGAGCGTCTTACAGGCTTC 3420  
3361 CAGAGGTTCCCGCTGCGCAGGACCATCTGAGAGTCTGGAGAGCGTCTTACAGGCTTC 3420  
3421 ACCACATAGACGCCCATTTCTTCCAGACTAAGCAGGACGAGGACAACTTCCCTTAC 3480  
3421 ACCACATAGACGCCCATTTCTTCCAGACTAAGCAGGACGAGGACAACTTCCCTTAC 3480  
3481 CTGTTAGCATACAGGCTACGTTGCGCCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3481 CTGTTAGCATACAGGCTACGTTGCGCCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3541 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTT 3600  
3541 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTT 3600  
3601 TATAGGCTGGGAGCGGTTCAAAACGAGGTTTACTACACACACCCCATTAACCAATACATC 3660  
3601 TATAGGCTGGGAGCGGTTCAAAACGAGGTTTACTACACACACCCCATTAACCAATACATC 3660  
3661 ATGGCATGATGTCCGCTGACCTGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720  
3661 ATGGCATGATGTCCGCTGACCTGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720  
3721 GTCTAGCAGTCTGGCGCGTATTTGCTGACCAAGCAGGAGTCTTATTCGCGGAGTTC 3780  
3721 GTCTAGCAGTCTGGCGCGTATTTGCTGACCAAGCAGGAGTCTTATTCGCGGAGTTC 3780  
3781 ATCATCTTGTCCGGAAGCGGCGCATATTCCCGACAGGAGTCTTATTCGCGGAGTTC 3840  
3781 ATCATCTTGTCCGGAAGCGGCGCATATTCCCGACAGGAGTCTTATTCGCGGAGTTC 3840  
3841 GATGATGGAAGTGTGGCGCTCACACCTCCCTTACATCGAAAGGGAATGCAGCTCGCC 3900

Db 3841 GATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGAACAGGAAATCCAGCTCGCC 3900  
Qy 3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCACCAACAGCAAGCGAGGCT 3960  
Db 3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCACCAACAGCAAGCGAGGCT 3960  
Qy 3961 GCTGCTCCGCTGGTGGAAATCAAGTGGCGGACCCCTCGAAGCCTTCTGCGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGGAAATCAAGTGGCGGACCCCTCGAAGCCTTCTGCGCGAAGCATATG 4020  
Qy 4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGCTTGTCCACTCTGCTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGCTTGTCCACTCTGCTGGCAACCC 4080  
Qy 4081 GCGATAGCATCACTGATGGCATTCACAGCCTCTATCACCAGCCGCTCACCACCCCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCATTCACAGCCTCTATCACCAGCCGCTCACCACCCCAACAT 4140  
Qy 4141 ACCCTCCTGTTTAAATCCTGGGGGATGGGTGGCGGCCCAACTTGTCTCCAGCGCT 4200  
Db 4141 ACCCTCCTGTTTAAATCCTGGGGGATGGGTGGCGGCCCAACTTGTCTCCAGCGCT 4200  
Qy 4201 GCTTCTGCTTTCGTAGGCGCGGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCGGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Qy 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGCGCTGGGCC 4320  
Db 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGCGCTGGGCC 4320  
Qy 4321 TTTAAGGTTCATGAGCGCGGAGTGCCTCCACCGAGGACCTGGTTAACTTACTCCCTGCT 4380  
Db 4321 TTTAAGGTTCATGAGCGCGGAGTGCCTCCACCGAGGACCTGGTTAACTTACTCCCTGCT 4380  
Qy 4381 ATCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTGGCAGCATACTGCTCGGCAC 4440  
Db 4381 ATCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTGGCAGCATACTGCTCGGCAC 4440  
Qy 4441 GTGGGCCAGGGAGGGGCTGTGAGTGAACCGGCTGATAGCGTTCGCTTCGCGG 4500  
Db 4441 GTGGGCCAGGGAGGGGCTGTGAGTGAACCGGCTGATAGCGTTCGCTTCGCGG 4500  
Qy 4501 GGTAAACAGGTTCCCGCAGCACTATGTGCTGAGAGGAGCGCTGACGAGCTCACT 4560  
Db 4501 GGTAAACAGGTTCCCGCAGCACTATGTGCTGAGAGGAGCGCTGACGAGCTCACT 4560  
Qy 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
Qy 4621 GAGGACTGTCCACGCCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Db 4621 GAGGACTGTCCACGCCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Qy 4681 ACGGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGCTCCTGCCGATTCGCGGA 4740  
Db 4681 ACGGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGCTCCTGCCGATTCGCGGA 4740  
Qy 4741 GTCCCTCTCTCTCATGTCAAAGTGGGTACAAGGAGTCTGGCGGGGACCGCATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCAAAGTGGGTACAAGGAGTCTGGCGGGGACCGCATCATG 4800  
Qy 4801 CAAACCACTGCTCCATGTGGAGCAGATCAACCGGACATGTGAANAACGGTTCCATGAG 4860  
Db 4801 CAAACCACTGCTCCATGTGGAGCAGATCAACCGGACATGTGAANAACGGTTCCATGAG 4860  
Qy 4861 ATCTGGGCGCTTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCAATTAACGGTAC 4920  
Db 4861 ATCTGGGCGCTTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCAATTAACGGTAC 4920  
Qy 4921 ACCACGGGCGCTTGACGCGCTCCCGCGGCCAAATTAATTCAGGCGCTGTGGGGGTG 4980

Db 4921 ACCACGGGCGCTTGACGCGCTCCCGCGCGCCAAATTAATTCAGGCGCTGTGGGGGTG 4980  
Qy 4981 GCTGCTCAGGAGTACGTGGAGGTTTACCGGGTGGGGATTTCCACTAGTTCAGCGGCAATG 5040  
Db 4981 GCTGCTCAGGAGTACGTGGAGGTTTACCGGGTGGGGATTTTCCACTAGTTCAGCGGCAATG 5040  
Qy 5041 ACCACTCACAACTGAAAGTGCCTGTGAGGTTCCGGCCCCCGAAATTTCTTCCAGAAAGTG 5100  
Db 5041 ACCACTCACAACTGAAAGTGCCTGTGAGGTTCCGGCCCCCGAAATTTCTTCCAGAAAGTG 5100  
Qy 5101 GATGGGTGGGGTGGACAGGTACGCTCCAGGTGCAAAACCCCTCTTACGGGAGGAGTGC 5160  
Db 5101 GATGGGTGGGGTGGACAGGTACGCTCCAGGTGCAAAACCCCTCTTACGGGAGGAGTGC 5160  
Qy 5161 ACATTCCTGTCGGGCTCAATCAATAGTGTGGGTGACAGCTCCCATGCGAGCGCGAA 5220  
Db 5161 ACATTCCTGTCGGGCTCAATCAATAGTGTGGGTGACAGCTCCCATGCGAGCGCGAA 5220  
Qy 5221 CCGGACGTAGCAGTGTCTCATCTCCATGCTCACCGACCCCTCCACATTTACGGCGAGACG 5280  
Db 5221 CCGGACGTAGCAGTGTCTCATCTCCATGCTCACCGACCCCTCCACATTTACGGCGAGACG 5280  
Qy 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Qy 5341 CTGTCTGGCGCTTCTTGAAGGCAATGCTACTCCCGTCTAGTCTCCCGGACGCTGAC 5400  
Db 5341 CTGTCTGGCGCTTCTTGAAGGCAATGCTACTCCCGTCTAGTCTCCCGGACGCTGAC 5400  
Qy 5401 CTCTCAGGCGCAACCTCTCTGCGGAGGATGGCGGGAAACATCACCCGCGTGAG 5460  
Db 5401 CTCTCAGGCGCAACCTCTCTGCGGAGGATGGCGGGAAACATCACCCGCGTGAG 5460  
Qy 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTTGGAGCGCTCCAGCGGAGGAGATGAG 5520  
Db 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTTGGAGCGCTCCAGCGGAGGAGATGAG 5520  
Qy 5521 AGGGAAGTATCGTTCGGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGACGATG 5580  
Db 5521 AGGGAAGTATCGTTCGGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGACGATG 5580  
Qy 5581 CCCATATGGGCAACCGGATTTACAACTTCCACTGTTAGAGTCTGGAAGAACCCGAC 5640  
Db 5581 CCCATATGGGCAACCGGATTTACAACTTCCACTGTTAGAGTCTGGAAGAACCCGAC 5640  
Qy 5641 TAGTCTCCCTCAGTGTGATACAGGTTGCTGCTGCGCTCCAGGCGCTCCGATACCA 5700  
Db 5641 TAGTCTCCCTCAGTGTGATACAGGTTGCTGCTGCGCTCCAGGCGCTCCGATACCA 5700  
Qy 5701 CCTCCAGGAGGAGGAGGAGGTTGCTGCTGCTCAGAAATCTTACCGTCTTCTGCTGGG 5760  
Db 5701 CCTCCAGGAGGAGGAGGAGGTTGCTGCTGCTCAGAAATCTTACCGTCTTCTGCTGGG 5760  
Qy 5761 GAGCTCGCCACAAAGACCTTCGGAGCTCCGAAATCGTGGCGCTGCAAGCGGCAACGGA 5820  
Db 5761 GAGCTCGCCACAAAGACCTTCGGAGCTCCGAAATCGTGGCGCTGCAAGCGGCAACGGA 5820  
Qy 5821 ACGGCTCTCTGACAGGCGCTCCGAGCAGCGGAGTCCGAGTCTCAGCGGACCGGCA 5880  
Db 5821 ACGGCTCTCTGACAGGCGCTCCGAGCAGCGGAGTCCGAGTCTCAGCGGACCGGCA 5880  
Qy 5881 TCCTCCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGCGACCGGCTTGG 5940  
Db 5881 TCCTCCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGCGACCGGCTTGG 5940  
Qy 5941 TCTACCGTAAGCGAGGAGCTAGTGAGGAGCTGCTGCTGCTGCTGCTGCTTACATGG 6000  
Db 5941 TCTACCGTAAGCGAGGAGCTAGTGAGGAGCTGCTGCTGCTGCTGCTGCTTACATGG 6000  
Qy 6001 ACAGGCGCTCATCAGCCATGCGTGGAGGAAACCAAGTGGCCATCAATGACATG 6060  
Db 6001 ACAGGCGCTCATCAGCCATGCGTGGAGGAAACCAAGTGGCCATCAATGACATG 6060

QY	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAAATCTCGAGCGCAAGC	6120
Db	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAAATCTCGAGCGCAAGC	6120
QY	6121	CTGGCGAGAAAGAGTCACTTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
Db	6121	CTGGCGAGAAAGAGTCACTTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
QY	6181	GTGCTCAGGAGATGAAGGGCAAGCGTCCACAGTTAAGCTAACTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAGGAGATGAAGGGCAAGCGTCCACAGTTAAGCTAACTTCTATCCGTGGAG	6240
QY	6241	GAGCCCTGTAAAGCTGACGCCCCACATTCGCGCAGATCTAAATTTGGCTATGGGGCAAG	6300
Db	6241	GAGCCCTGTAAAGCTGACGCCCCACATTCGCGCAGATCTAAATTTGGCTATGGGGCAAG	6300
QY	6301	GAGCTCCGGAACCTATCCAGCAAGCGCGTTAAACACATCCGCTCCGCTGCGAAGGACTTG	6360
Db	6301	GAGCTCCGGAACCTATCCAGCAAGCGCGTTAAACACATCCGCTCCGCTGCGAAGGACTTG	6360
QY	6361	CTGGAAGACACTGAGACACCAATTGACACCAATCATGTCGCAAAATGAGGTTTCTGCG	6420
Db	6361	CTGGAAGACACTGAGACACCAATTGACACCAATCATGTCGCAAAATGAGGTTTCTGCG	6420
QY	6421	GTCACACGAGAGAGGGGCGCCAGCTGCGCTTATCGTATTCACAGATTTGGGG	6480
Db	6421	GTCACACGAGAGAGGGGCGCCAGCTGCGCTTATCGTATTCACAGATTTGGGG	6480
QY	6481	GTTCTGTGTGCGAGAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
Db	6481	GTTCTGTGTGCGAGAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
QY	6541	ATGGGCTCTTCAACGATTCGAATCTCTCTGGAAGCGGCTCGAGTTCTCTGGTGAAT	6600
Db	6541	ATGGGCTCTTCAACGATTCGAATCTCTCTGGAAGCGGCTCGAGTTCTCTGGTGAAT	6600
QY	6601	GCTCTGAAGCGAAGAAATGSCCTATGCGCTTCGATATGACACCCGCTGTTGACTCA	6660
Db	6601	GCTCTGAAGCGAAGAAATGSCCTATGCGCTTCGATATGACACCCGCTGTTGACTCA	6660
QY	6661	ACGGTCACTGAGAAATGACATCCGCTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
Db	6661	ACGGTCACTGAGAAATGACATCCGCTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
QY	6721	CCGGAAGCCAGACAGGCCAATAAGGTCGCTCAGAGCGGCTTTACATCGGGGCCCCCTG	6780
Db	6721	CCGGAAGCCAGACAGGCCAATAAGGTCGCTCAGAGCGGCTTTACATCGGGGCCCCCTG	6780
QY	6781	ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTCGCGCGAGCGGTGTACTGACG	6840
Db	6781	ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTCGCGCGAGCGGTGTACTGACG	6840
QY	6841	ACCAGCTCGGTAATACCTTCATGTTACTTGAAGCGCGCTCGCGCTGTCCAGCTGCG	6900
Db	6841	ACCAGCTCGGTAATACCTTCATGTTACTTGAAGCGCGCTCGCGCTGTCCAGCTGCG	6900
QY	6901	AGCTCCAGGACTGCGACGATGCTCGTATCGGAGACGACCTTGTCTGTGTAAGAGC	6960
Db	6901	AGCTCCAGGACTGCGACGATGCTCGTATCGGAGACGACCTTGTCTGTGTAAGAGC	6960
QY	6961	GCGGGACCCAGAGACGAGCGGCTTACGGGCTTCCAGGAGCTATGACTAGATAC	7020
Db	6961	GCGGGACCCAGAGACGAGCGGCTTACGGGCTTCCAGGAGCTATGACTAGATAC	7020
QY	7021	TCTGCCCTCCCTGGGACCGCCCAACAGATACGACTTGGAGTTGATACATATGC	7080
Db	7021	TCTGCCCTCCCTGGGACCGCCCAACAGATACGACTTGGAGTTGATACATATGC	7080
QY	7081	TCTTCAATGTGTCACTGCGGACGATGATCTGGCAAAAGGCTGTACTATCTCACCCGT	7140
Db	7081	TCTTCAATGTGTCACTGCGGACGATGATCTGGCAAAAGGCTGTACTATCTCACCCGT	7140

QY	7141	GACCCCAACACCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT	7200
Db	7141	GACCCCAACACCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT	7200
QY	7201	TCTGGCTAGGCAACATCATGTATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
Db	7201	TCTGGCTAGGCAACATCATGTATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
QY	7261	ACTCATTTCTTCCATCTTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTGAG	7320
Db	7261	ACTCATTTCTTCCATCTTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTGAG	7320
QY	7321	ATCTAGGGGCTTGTACTCCATTTGAGGCACTTGAACCTACTCAGATCATTTCAACGACTC	7380
Db	7321	ATCTAGGGGCTTGTACTCCATTTGAGGCACTTGAACCTACTCAGATCATTTCAACGACTC	7380
QY	7381	CATGGCTTAGCGCATTTTTCACCTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
Db	7381	CATGGCTTAGCGCATTTTTCACCTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
QY	7441	TCAATGCTCAGGAACCTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGGCCAGAACT	7500
Db	7441	TCAATGCTCAGGAACCTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGGCCAGAACT	7500
QY	7501	GTCGGCTTAGGCTACTGTCCTCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
Db	7501	GTCGGCTTAGGCTACTGTCCTCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
QY	7561	AACCTGGGAGTAGGACCAAGCTCAAACTCACCTCCATCCCGCTGCTCCAGTTGGAT	7620
Db	7561	AACCTGGGAGTAGGACCAAGCTCAAACTCACCTCCATCCCGCTGCTCCAGTTGGAT	7620
QY	7621	TATTCAGCTGCTTCTGTTTACAGGGGGAGACATATATCACAGCTGTCTCGT	7680
Db	7621	TATTCAGCTGCTTCTGTTTACAGGGGGAGACATATATCACAGCTGTCTCGT	7680
QY	7681	GCCGACCCGCTGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT	7740
Db	7681	GCCGACCCGCTGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT	7740
QY	7741	CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT	7800
Db	7741	CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT	7800
QY	7801	TTTCCCTTT	7860
Db	7801	TTTCCCTTT	7860
QY	7861	TTTTTCTCTTT	7920
Db	7861	TTTTTCTCTTT	7920
QY	7921	TAGCTGTGAAGGCTCGTCCGCTGAGCGCTTGAAGTGTGCTGATCTGCTGCTCTGCTC	7980
Db	7921	TAGCTGTGAAGGCTCGTCCGCTGAGCGCTTGAAGTGTGCTGATCTGCTGCTCTGCTC	7980
QY	7981	AGATCAAGT 7989	
Db	7981	AGATCAAGT 7989	

## RESULT 2

ADP86264

ID ADP86264 standard; DNA; 11313 Bp.

XX AC ADP86264;

XX AC ADP86264;

DT 23-SEP-2004 (first entry)

DE Hepatitis C virus Con-1 replicon I377/NS3-3', plasmid DNA.

XX Hepatitis C virus; HCV; anti-HCV agent; HCV infection; therapy; plasmid;

KW ds.



Db 1621 TGGCTCTCCTCAAGCGTATTCAACAAAGGGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680  
QY 1681 ATGGGATCTGATCTGGGGCCCTGGTGACATGCTTTTACATGTGTTAGTCGAGGTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCCCTGGTGACATGCTTTTACATGTGTTAGTCGAGGTAAAA 1740  
QY 1741 AACCTCTAGGCCCCCGGAAACACAGGGGACGTGGTTTTTCTTTTGAAGAACACGATAATACC 1800  
Db 1741 AACCTCTAGGCCCCCGGAAACACAGGGGACGTGGTTTTTCTTTTGAAGAACACGATAATACC 1800  
QY 1801 ATGGCGCCTATTAGCGCTACTCCCAACAGACGGAGGCCCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGCGCCTATTAGCGCTACTCCCAACAGACGGAGGCCCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCTCACAGCGCGGGACAGGAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTCACAGCGCGGGACAGGAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACAAATCTTTTCTGGCGACTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 1980  
Db 1921 ACACAAATCTTTTCTGGCGACTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 1980  
QY 1981 GGCTCAAGACCTTTGCGGCCCAAGAGGCCCAATACCAAAATGTACACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCTTTGCGGCCCAAGAGGCCCAATACCAAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACTCTCGGCTGGCAAGCGGCCCGCGGGCGGCTTCCTTTGACACCAATGCACCTGC 2100  
Db 2041 CAGGACTCTCGGCTGGCAAGCGGCCCGCGGGCGGCTTCCTTTGACACCAATGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGGTCACGAGGATGCGCATGTCTCATCTCCGTGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGGTCACGAGGATGCGCATGTCTCATCTCCGTGCGCGCGG 2160  
QY 2161 GGCAGCAGAGGGGAGCCTACTCTCCCGAGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
Db 2161 GGCAGCAGAGGGGAGCCTACTCTCCCGAGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGCTCCACTGCTCTGCCCCCTCGGGCAGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
Db 2221 GCGGCTCCACTGCTCTGCCCCCTCGGGCAGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCCGAGGGGTTCCGAAGGCGGTGGACTTTGTATACCGTCCAGCTATGGAAACCACTATG 2340  
Db 2281 ACCCGAGGGGTTCCGAAGGCGGTGGACTTTGTATACCGTCCAGCTATGGAAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTCCAGGACAACTCGTCCCTCCGCGCGTACCGCAGACATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTCCAGGACAACTCGTCCCTCCGCGCGTACCGCAGACATTTCCAGGTG 2400  
QY 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGCACTGTTTGGTGGTGGTGGTGGTGGTGGT 2460  
Db 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGCACTGTTTGGTGGTGGTGGTGGTGGTGGT 2460  
QY 2461 GCCAAGGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCCGCCCACTTAGTTCGGG 2520  
Db 2461 GCCAAGGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCCGCCCACTTAGTTCGGG 2520  
QY 2521 GGGTATATGTGAAGCAATGTATGACATCAATATGTATGAGTGGCAAGTTCCTGCGACGGTGGT 2580  
Db 2521 GGGTATATGTGAAGCAATGTATGACATCAATATGTATGAGTGGCAAGTTCCTGCGACGGTGGT 2580  
QY 2581 ACCAGGGTGGCCCATACAGTACTCCACTATGGCAAGTTCCTGCGACGGTGGTGGTGGTGG 2640  
Db 2581 ACCAGGGTGGCCCATACAGTACTCCACTATGGCAAGTTCCTGCGACGGTGGTGGTGGTGG 2640  
QY 2641 TCTGGGGGCGCTTATGACATCAATATGTATGAGTGGCAAGTTCCTGCGACGGTGGTGGTGG 2700  
Db 2641 TCTGGGGGCGCTTATGACATCAATATGTATGAGTGGCAAGTTCCTGCGACGGTGGTGGTGG 2700  
QY 2701 ATCTGGGATCGGCACTGCTTGGACCAAGCGGAGACGGCTGGAGCGGACTCTGCTGCTG 2760  
Db 2701 ATCTGGGATCGGCACTGCTTGGACCAAGCGGAGACGGCTGGAGCGGACTCTGCTGCTG 2760

QY 2761 CTGCCACCGCTAGCGCTCCGGATCGGTACCGTGCACATCCAAATCCAAATCCAAATCCAAATCC 2820  
Db 2761 CTGCCACCGCTAGCGCTCCGGATCGGTACCGTGCACATCCAAATCCAAATCCAAATCCAAATCC 2820  
QY 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCAATCCCAATCCCAATCC 2880  
Db 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCAATCCCAATCCCAATCC 2880  
QY 2881 AAGGGGCGGAGGCACTTCTTCTGCTCAATTTCCAAAGAAAGAAATGTAGTGCCTGCGCG 2940  
Db 2881 AAGGGGCGGAGGCACTTCTTCTGCTCAATTTTCCAAAGAAAGAAATGTAGTGCCTGCGCG 2940  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTTACCGGGGCCCTTGTATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTTACCGGGGCCCTTGTATCCGTC 3000  
QY 3001 ATACCACTAGCGGAGAGCTTCTGCTAGCAACCGACGCTCTAATGACGGGCTTTACC 3060  
Db 3001 ATACCACTAGCGGAGAGCTTCTGCTAGCAACCGACGCTCTAATGACGGGCTTTACC 3060  
QY 3061 GGGCAATTTCCAGCTCAGTGTATCGACTGCAATACATGTGTCAACCGACAGTGTGACTCAGC 3120  
Db 3061 GGGCAATTTCCAGCTCAGTGTATCGACTGCAATACATGTGTCAACCGACAGTGTGACTCAGC 3120  
QY 3121 CTGGACCGGACCTTCCATTTGAGACGACGACCGGTGCCAAGACGCGGTGTCAACGCTG 3180  
Db 3121 CTGGACCGGACCTTCCATTTGAGACGACGACCGGTGCCAAGACGCGGTGTCAACGCTG 3180  
QY 3181 CAGGGGCGGAGGAGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3181 CAGGGGCGGAGGAGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGCGCTTCGGGCACTTCTCGATTCCTCGTCTGTGCGAGTGTATGACGGCGGTGT 3300  
Db 3241 GAACGCGCTTCGGGCACTTCTCGATTCCTCGTCTGTGCGAGTGTATGACGGCGGTGT 3300  
QY 3301 GCTTGTAGCAGCTACGCGCCCGGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTAGCAGCTACGCGCCCGGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
QY 3361 CAGGGTTCGGCTCTGCGAGCACTCTGGAGTTCTGGAGAGCGTCTTTACAGGCTC 3420  
Db 3361 CAGGGTTCGGCTCTGCGAGCACTCTGGAGTTCTGGAGAGCGTCTTTACAGGCTC 3420  
QY 3421 ACCCATAGAGCCCACTTTCTGCCAGACTTAAGCAGCAGGAGACACTTTCCCTAC 3480  
Db 3421 ACCCATAGAGCCCACTTTCTGCCAGACTTAAGCAGCAGGAGACACTTTCCCTAC 3480  
QY 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATAGGCTAAAGCCTACGCTGCAAGGGCAAGCCCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATAGGCTAAAGCCTACGCTGCAAGGGCAAGCCCTGCTG 3600  
QY 3601 TATAGCTCGGAGCGGTTCAAAACAGGTTTACTACACACACCCCATACCAATACATC 3660  
Db 3601 TATAGCTCGGAGCGGTTCAAAACAGGTTTACTACACACACCCCATACCAATACATC 3660  
QY 3661 ATGGCATGATGCTGGCTGACCTGGAGTGTGTCAGAGCACTGGGTGTGTAGGCGA 3720  
Db 3661 ATGGCATGATGCTGGCTGACCTGGAGTGTGTCAGAGCACTGGGTGTGTAGGCGA 3720  
QY 3721 GTCTAGCAGCTCTGGCGCGTATTTGCTGTGACAAAGCAGCGTGGTCAATTTGGGCGAG 3780  
Db 3721 GTCTAGCAGCTCTGGCGCGTATTTGCTGTGACAAAGCAGCGTGGTCAATTTGGGCGAG 3780  
QY 3781 ATCTCTTGTCCGGAAGCGGCCCATCTTCCGACAGGAAAGTCTTTTACCGGGAGTTC 3840  
Db 3781 ATCTCTTGTCCGGAAGCGGCCCATCTTCCGACAGGAAAGTCTTTTACCGGGAGTTC 3840



3841 GATGAGATGGAAGTGGGCTCACACTCCCTTACATCGAACAGGGAATGACGTGCGC 3900  
3841 GATGAGATGGAAGTGGGCTCACACTCCCTTACATCGAACAGGGAATGACGTGCGC 3900  
3901 GAACAAATCAAAAGAGGCAATCGGTTGCTGCAAAAGGCAACCAAGAGCGGAGGCT 3960  
3901 GAACAAATCAAAAGAGGCAATCGGTTGCTGCAAAAGGCAACCAAGAGCGGAGGCT 3960  
3961 GCTGCTCCGTTGGAATCAAGTGGCGGACCCCTCGAAGCCTTCTGGCGAAGCATATG 4020  
3961 GCTGCTCCGTTGGAATCAAGTGGCGGACCCCTCGAAGCCTTCTGGCGAAGCATATG 4020  
4021 TGGAAATTCATCAGCGGATCAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCC 4080  
4021 TGGAAATTCATCAGCGGATCAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCC 4080  
4081 CGCATAGCATCACTGATGGCAATTCACAGCTTATCACCAGCCCGCTCACCACCCAAAT 4140  
4081 CGCATAGCATCACTGATGGCAATTCACAGCTTATCACCAGCCCGCTCACCACCCAAAT 4140  
4141 ACCCTCTGTTTAAATCTGAGGCGGATGGGTGGCGGCCAACCTTGTCTCTCCACGCT 4200  
4141 ACCCTCTGTTTAAATCTGAGGCGGATGGGTGGCGGCCAACCTTGTCTCTCCACGCT 4200  
4201 GCTTCTGCTTTGATAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCGCTGGG 4260  
4201 GCTTCTGCTTTGATAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCGCTGGG 4260  
4261 AAGTGTCTTGTGGAATTTTGGCAGTTATGAGCAGGCGGTGGCAGGCGCGCTCGTGCC 4320  
4261 AAGTGTCTTGTGGAATTTTGGCAGTTATGAGCAGGCGGTGGCAGGCGCGCTCGTGCC 4320  
4321 TTTAAGGTCAAGCGGCGAGATGCCCTCCACCGAGGACTGGTTAACTACTCCCTGCT 4380  
4321 TTTAAGGTCAAGCGGCGAGATGCCCTCCACCGAGGACTGGTTAACTACTCCCTGCT 4380  
4381 ATCCTCTCCCTGCGGCGCTAGTCTCGGGTCTGTCGCGAGCATACTCGTCGGCAC 4440  
4381 ATCCTCTCCCTGCGGCGCTAGTCTCGGGTCTGTCGCGAGCATACTCGTCGGCAC 4440  
4441 GTGGCGCCAGGAGGAGGCGCTGTGAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
4441 GTGGCGCCAGGAGGAGGCGCTGTGAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
4501 GGTAAACAGCTCTCCCGGCGAGTATGTCGCTGAGCGAGCGCTGCGACGCTGTCTACT 4560  
4501 GGTAAACAGCTCTCCCGGCGAGTATGTCGCTGAGCGAGCGCTGCGACGCTGTCTACT 4560  
4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCCAGTGGATCAAC 4620  
4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCCAGTGGATCAAC 4620  
4621 GAGGACTGTCTCCAGCGCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGATGATGC 4680  
4621 GAGGACTGTCTCCAGCGCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGATGATGC 4680  
4681 ACGTGTTCAGTATTTAAGACTGCTGCTCAGTCCAGTCTGCGCGGATGCGCGGA 4740  
4681 ACGTGTTCAGTATTTAAGACTGCTGCTCAGTCCAGTCTGCGCGGATGCGCGGA 4740  
4741 GTCCCTCTCTCATGTCAAAGTGGTACAGGAGTCTGGCGGGCGACGCAATCATG 4800  
4741 GTCCCTCTCTCATGTCAAAGTGGTACAGGAGTCTGGCGGGCGACGCAATCATG 4800  
4801 CAAACCACTGCCCATGTGGAGCACAGATCAACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
4801 CAAACCACTGCCCATGTGGAGCACAGATCAACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
4861 ATCGTGGGCGTAGGACTGTAGTAAACAGTGGCATGGAACATTTCCCGATTACGCGTAC 4920  
4861 ATCGTGGGCGTAGGACTGTAGTAAACAGTGGCATGGAACATTTCCCGATTACGCGTAC 4920  
4921 ACCACGGGCGCTGACGCGCTCCCGCGGCCAAATTAATTTAGGGCGCTGTGGCGGGTG 4980

4921 ACCACGGGCGCTGACGCGCTCCCGCGGCCAAATTAATTTAGGGCGCTGTGGCGGGTG 4980  
4981 GCTGCTGAGGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAGTACGCGGATG 5040  
4981 GCTGCTGAGGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAGTACGCGGATG 5040  
5041 ACCACTGACAACTGAAAGTCCCGTGTACAGTTCCGGCCCCCGAATTTCTTACAGAGTG 5100  
5041 ACCACTGACAACTGAAAGTCCCGTGTACAGTTCCGGCCCCCGAATTTCTTACAGAGTG 5100  
5101 GATGGGTGCGGTTGCCACAGGTACGCTCCAGCTGCAAAACCCCTCTTACGGAGGAGTGC 5160  
5101 GATGGGTGCGGTTGCCACAGGTACGCTCCAGCTGCAAAACCCCTCTTACGGAGGAGTGC 5160  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCGAA 5220  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCGAA 5220  
5221 CCGGAGGTAGCAGTGTCACTTCCATGTCAACGACCCCTCCCATTTACGCGGAGACG 5280  
5221 CCGGAGGTAGCAGTGTCACTTCCATGTCAACGACCCCTCCCATTTACGCGGAGACG 5280  
5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGGCCAGCTCATAGCTAGCCAG 5340  
5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGGCCAGCTCATAGCTAGCCAG 5340  
5341 CTGTCTGGCGCTTCTTGAAGCAACATGCACTACCCGCTCATGCTCCCGGACGCTGAC 5400  
5341 CTGTCTGGCGCTTCTTGAAGCAACATGCACTACCCGCTCATGCTCCCGGACGCTGAC 5400  
5401 CTCATCGAGGCCAACCTCTGTGGCGGAGAGATGGGCGGAAACATCACCCGCTGGAG 5460  
5401 CTCATCGAGGCCAACCTCTGTGGCGGAGAGATGGGCGGAAACATCACCCGCTGGAG 5460  
5461 TCAGAAAAATAGGTAGTAATTTTGACTCTTTTCGAGCGCTCCAAGCGGAGGAGTACG 5520  
5461 TCAGAAAAATAGGTAGTAATTTTGACTCTTTTCGAGCGCTCCAAGCGGAGGAGTACG 5520  
5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGAAATTCCTCGAGCGATG 5580  
5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGAAATTCCTCGAGCGATG 5580  
5581 CCATATCGGACGCGCGATTTACAACTCTTACCTGTTAGTCTCTGGAAGACCCGAC 5640  
5581 CCATATCGGACGCGCGATTTACAACTCTTACCTGTTAGTCTCTGGAAGACCCGAC 5640  
5641 TACGTCCTCTCCAGTGTACACGGGTGTCAATTCGCGCTGCAAGCCCTCCGATACCA 5700  
5641 TACGTCCTCTCCAGTGTACACGGGTGTCAATTCGCGCTGCAAGCCCTCCGATACCA 5700  
5701 CTTCCAGGAGGAGGAGGAGTCTGCTGTCAGATCTACCGTGTCTTCTGCTTGGG 5760  
5701 CTTCCAGGAGGAGGAGGAGTCTGCTGTCAGATCTACCGTGTCTTCTGCTTGGG 5760  
5761 GAGTCTCGCCACAAAGACCTTCGCGAGTCCGAAATCTGCGCGCTCGACGCGGACGCA 5820  
5761 GAGTCTCGCCACAAAGACCTTCGCGAGTCCGAAATCTGCGCGCTCGACGCGGACGCA 5820  
5821 ACGGCTCTCTGACACAGCCCTCCAGCGAGCGGAGTCCGATCTGAGTTCGATG 5880  
5821 ACGGCTCTCTGACACAGCCCTCCAGCGAGCGGAGTCCGATCTGAGTTCGATG 5880  
5881 TCTTCATGCCCCCTTTCGAGGGGAGCCGGGGATCCCGATCTCAGCGACGCGTCTTG 5940  
5881 TCTTCATGCCCCCTTTCGAGGGGAGCCGGGGATCCCGATCTCAGCGACGSGTCTTG 5940  
5941 TCTACCGTAAGCGAGGAGGCTAGTGGAGCGTCTGCTGCTCGATGCTTACACATGG 6000  
5941 TCTACCGTAAGCGAGGAGGCTAGTGGAGCGTCTGCTGCTCGATGCTTACACATGG 6000  
6001 ACGGCGCCCTGATCAGGCCATGCGCTGGGAGGAAACCAAGTGGCCATCAATGACATG 6060

Db 6001 ACAGGGCCCTGATCAGCCATGCGCTCGGAGAAACCAAGCTGCCCATCATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCGGTCAACCAAACTTGGTCTATGCTACAAATCTCCGACGCGAAGC 6120  
Db 6061 AGCAACTCTTTGCTCGGTCAACCAAACTTGGTCTATGCTACAAATCTCCGACGCGAAGC 6120  
QY 6121 CTGCGGACAGAGAGTCACTTTGACAGACTGACAGTCTGAGACGACCACTACCGGAC 6180  
Db 6121 CTGCGGACAGAGAGTCACTTTGACAGACTGACAGTCTGAGACGACCACTACCGGAC 6180  
QY 6181 GTGCTCAAGAGATGAAGCGAAGGCTCCACAGTTTAAAGCTTAACTTCTATCCGTTGGAG 6240  
Db 6181 GTGCTCAAGAGATGAAGCGAAGGCTCCACAGTTTAAAGCTTAACTTCTATCCGTTGGAG 6240  
QY 6241 GAAGCTGTAAAGTGAAGCGCCCAATTCGCGAGATCTAAATTTGGCTATGCGGCAAAAG 6300  
Db 6241 GAAGCTGTAAAGTGAAGCGCCCAATTCGCGAGATCTAAATTTGGCTATGCGGCAAAAG 6300  
QY 6301 GAGTCCGGAACCTATCCAGCAGGCGCTTAAACCATCGCTCCGTTGGAGGACTTG 6360  
Db 6301 GAGTCCGGAACCTATCCAGCAGGCGCTTAAACCATCGCTCCGTTGGAGGACTTG 6360  
QY 6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATGAGTTTCTGC 6420  
Db 6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATGAGTTTCTGC 6420  
QY 6421 GTCCAAACAGAGAGAGGCGCGCAAGCAGTCTGCTTATCGTATTCACAGATTTGGGG 6480  
Db 6421 GTCCAAACAGAGAGAGGCGCGCAAGCAGTCTGCTTATCGTATTCACAGATTTGGGG 6480  
QY 6481 GTTCGTGTGTCGAGAAATGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG 6540  
Db 6481 GTTCGTGTGTCGAGAAATGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG 6540  
QY 6541 ATGGGCTCTTCATACGATTCCAATACTCTCTCGACAGCGGTTCGAGTTCTCTGGTGAAT 6600  
Db 6541 ATGGGCTCTTCATACGATTCCAATACTCTCTCGACAGCGGTTCGAGTTCTCTGGTGAAT 6600  
QY 6601 GCCTGAAAGCAGAGAAATGCCCTATGCGCTTCGCATATGACACCGCTGTTTGTACTCA 6660  
Db 6601 GCCTGAAAGCAGAGAAATGCCCTATGCGCTTCGCATATGACACCGCTGTTTGTACTCA 6660  
QY 6661 ACAGTCACTGAGATGACATCGTGTGAGGAGTCAATCTCAATGTTGTGACTGGCC 6720  
Db 6661 ACAGTCACTGAGATGACATCGTGTGAGGAGTCAATCTCAATGTTGTGACTGGCC 6720  
QY 6721 CCGGAAGCCAGACAGGCCATAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTG 6780  
Db 6721 CCGGAAGCCAGACAGGCCATAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTG 6780  
QY 6781 ACTAATCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGACTGACG 6840  
Db 6781 ACTAATCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGACTGACG 6840  
QY 6841 ACCAGTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTGCGGCTGTCGAGCTGG 6900  
Db 6841 ACCAGTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTGCGGCTGTCGAGCTGG 6900  
QY 6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
Db 6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
QY 6961 GCGGGACCCAGAGAGCAGGCGGAGCTACGCGCTTACCGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAGAGAGCAGGCGGAGCTACGCGCTTACCGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTTGGGACCGGCCCAACCAAGATACGACTTGGAGTTGATACATCATGC 7080  
Db 7021 TCTGCCCCCTTGGGACCGGCCCAACCAAGATACGACTTGGAGTTGATACATCATGC 7080  
QY 7081 TCTCCAAATGTGTCAGTCCGCGACGATGATCTGCGCAAAAGGCTGACTATCTCACCCCT 7140  
Db 7081 TCTCCAAATGTGTCAGTCCGCGACGATGATCTGCGCAAAAGGCTGACTATCTCACCCCT 7140

QY 7141 GACCCCAACACCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCCAACACCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT 7200  
QY 7201 TCCTGCTAGGCAACATCATCATGATATGCCCCACCTTGTGGCAAGGATGATCCCTGATG 7260  
Db 7201 TCCTGCTAGGCAACATCATCATGATATGCCCCACCTTGTGGCAAGGATGATCCCTGATG 7260  
QY 7261 ACTCATTTTCTTCCATCCTTCTAGCTCAGGAACAACCTTGAAGGACCCCTAGATTCTCAG 7320  
Db 7261 ACTCATTTTCTTCCATCCTTCTAGCTCAGGAACAACCTTGAAGGACCCCTAGATTCTCAG 7320  
QY 7321 ATCTAGGGGCTGTTTATCTCATTTAGGCCACTTGACCTACCTCAGATCATTAAGACACTC 7380  
Db 7321 ATCTAGGGGCTGTTTATCTCATTTAGGCCACTTGACCTACCTCAGATCATTAAGACACTC 7380  
QY 7381 CATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
Db 7381 CATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
QY 7441 TCATGCTCAGGAACCTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGCGCAAGT 7500  
Db 7441 TCATGCTCAGGAACCTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGCGCAAGT 7500  
QY 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGTGCACCTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGTGCACCTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAACTCCGCGCTCGTCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAACTCCGCGCTCGTCCAGTTGGAT 7620  
QY 7621 TTATCCAGCTGCTTCTGTTGCTGTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680  
Db 7621 TTATCCAGCTGCTTCTGTTGCTGTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680  
QY 7681 GCCGACCCCGCTGTTTCTGTTGCTGCTTACTTCTTCTGTTAGGGGTAGGCTCTAT 7740  
Db 7681 GCCGACCCCGCTGTTTCTGTTGCTGCTTACTTCTTCTGTTAGGGGTAGGCTCTAT 7740  
QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
QY 7801 TTTCCCTTT 7860  
Db 7801 TTTCCCTTT 7860  
QY 7861 TTTTTCCTTT 7920  
Db 7861 TTTTTCCTTT 7920  
QY 7921 TAGCTGTGAAGGTCGCTGAGCGCTTGTACTCAGAGAGTGTGATCTAGTGGCTCTCTGC 7980  
Db 7921 TAGCTGTGAAGGTCGCTGAGCGCTTGTACTCAGAGAGTGTGATCTAGTGGCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

RESULT 3  
AAA98968  
ID AAA98968 standard; DNA; 7989 BP.  
XX  
AC AAA98968;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Hepatitis C virus DNA fragment SEQ ID NO: 4.  
XX  
KW Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.

XX Hepatitis C virus.  
 OS DE19915178-A1.  
 XX 05-OCT-2000.  
 XX 03-APR-1999; 99DE-01015178.  
 XX 03-APR-1999; 99DE-01015178.  
 XX (UYMA-) UNIV MAINZ GUTENBERG JOHANNES.  
 XX Bartenschlager R;  
 XX WPI; 2000-629140/61.  
 XX Cell culture system for hepatitis C virus, useful e.g. in screening for  
 PT therapeutic agents, comprises human hepatoma cells containing a viral RNA  
 PT construct that includes a selectable gene.  
 XX  
 XX Claim 8; Page 37-43; 58pp; German.  
 XX This invention describes a novel Hepatitis C virus (HCV) cell culture  
 CC system comprising human hepatoma cells that contain an integrated HCV-RNA  
 CC construct (I). (i) contains the HCV-specific RNA segments 5'-NTR (non-  
 CC translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-  
 CC NTR, and a selectable (marker) gene (II). The cell cultures, and/or (i),  
 CC are used to prepare, evaluate and/or test therapeutic and/or diagnostic  
 CC agents for HCV infections, and to prepare vaccines against HCV infection  
 CC (particularly preparation of attenuated HCV). The can also be used for  
 CC preparation of a liver-specific delivery system for gene therapy, and to  
 CC identify cells permissive for HCV replication. Virus RNA replicates  
 CC autonomously and with high efficiency in this cellular system, so that  
 CC variations in replication rates can be measured (for screening antiviral  
 CC agents) quantitatively or qualitatively, using standard laboratory  
 CC equipment. Efficient replication of HCV RNA is only achieved when the  
 CC specified RNA segments are present and when the transfected cells are  
 CC maintained under permanent selection pressure  
 XX  
 XX Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 7987.4; DB 3; Length 7989;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCCAGCCCCCATGGGGGGGACACTCCACATAGATCACTCCCTGTGAGGAAGTACTG 60  
 DB 1 GCCAGCCCCCATGGGGGGGACACTCCACATAGATCACTCCCTGTGAGGAAGTACTG 60  
 QY 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTGTGCTGCGAGCTCCAGGAC 120  
 DB 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTGTGCTGCGAGCTCCAGGAC 120  
 QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCGAG 180  
 DB 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCGAG 180  
 QY 181 GACACCGGTCTCTTCTTGATCAACCCGCTCAATGCTGAGATTTGGCGTGCCTCC 240  
 DB 181 GACACCGGTCTCTTCTTGATCAACCCGCTCAATGCTGAGATTTGGCGTGCCTCC 240  
 QY 241 GCGAGATCTGTAGCCGAGTAGTGTGGTCCGGAAGGCCCTTGTGTTACTCCCTGATAGG 300  
 DB 241 GCGAGATCTGTAGCCGAGTAGTGTGGTCCGGAAGGCCCTTGTGTTACTCCCTGATAGG 300  
 QY 301 GTGCTTGGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAAGATCCTAAAC 360  
 DB 301 GTGCTTGGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAAGATCCTAAAC 360  
 QY 361 CTCGAAGAAACCAAGGGCGCCATGATTGACAGATGATTCGACCGAGTTCTC 420  
 DB 361 CTCGAAGAAACCAAGGGCGCCATGATTGACAGATGATTCGACCGAGTTCTC 420

QY 421 CGGCGCTTGGTGGAGAGGCTATTCCGCTATGCTGGGCACACAGCAATCGGCTGCT 480  
 DB 421 CGGCGCTTGGTGGAGAGGCTATTCCGCTATGCTGGGCACACAGCAATCGGCTGCT 480  
 QY 481 CTGATGCCGCGCTGTTCCGCTGTGAGCGAGGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
 DB 481 CTGATGCCGCGCTGTTCCGCTGTGAGCGAGGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
 QY 541 ACCTGTCCGCTGCTGTAATGAACTGAGACGAGGAGCGCGGCTATGCTGGTGGGCA 600  
 DB 541 ACCTGTCCGCTGCTGTAATGAACTGAGACGAGGAGCGCGGCTATGCTGGTGGGCA 600  
 QY 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
 DB 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
 QY 661 TGCTATTGGGCGAAGTGCAGGAGGATCTCCTGTCTATCTCACCTTCTCTGCCGAGA 720  
 DB 661 TGCTATTGGGCGAAGTGCAGGAGGATCTCCTGTCTATCTCACCTTCTCTGCCGAGA 720  
 QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGCTACTGCC 780  
 DB 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGCTACTGCC 780  
 QY 781 CATTTCGACCCACCAAGCGAATCATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
 DB 781 CATTTCGACCCACCAAGCGAATCATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
 QY 841 TTGTCGATCAGGATGATCTGGAAGAGCATAGGGGCTCGCGCCAGCCGAACTGTTCG 900  
 DB 841 TTGTCGATCAGGATGATCTGGAAGAGCATAGGGGCTCGCGCCAGCCGAACTGTTCG 900  
 QY 901 CCAGGCTCAAGGCGCGATGCGCGAGGAGATCTGCTGTGACCATCGCGATGCT 960  
 DB 901 CCAGGCTCAAGGCGCGATGCGCGAGGAGATCTGCTGTGACCATCGCGATGCT 960  
 QY 961 GCTTCCGGAATCATGCTGGAATGCGCGCTTTCTTGGATTGATGACGTTGCGCGG 1020  
 DB 961 GCTTCCGGAATCATGCTGGAATGCGCGCTTTCTTGGATTGATGACGTTGCGCGG 1020  
 QY 1021 TGGTGTGGGCGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTCCTGAAGAGC 1080  
 DB 1021 TGGTGTGGGCGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTCCTGAAGAGC 1080  
 QY 1081 TTGCGGCGAATGCGCTGACCGCTTCTGCTGCTTTAGGTTATCGCGCTCCGATTCGC 1140  
 DB 1081 TTGCGGCGAATGCGCTGACCGCTTCTGCTGCTTTAGGTTATCGCGCTCCGATTCGC 1140  
 QY 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGAGTTCTTCTGAGTTTAAACAGACACACAG 1200  
 DB 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGAGTTCTTCTGAGTTTAAACAGACACACAG 1200  
 QY 1201 GTTTCCTCTAGCGGATCAATCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGCG 1260  
 DB 1201 GTTTCCTCTAGCGGATCAATCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGCG 1260  
 QY 1261 CGAAGCGCTTGGAAATGAGCGGCTGCGTTTCTGCTATATGTTATTTCCACATATTG 1320  
 DB 1261 CGAAGCGCTTGGAAATGAGCGGCTGCGTTTCTGCTATATGTTATTTCCACATATTG 1320  
 QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCGCTGTCTTCTTGAAGCATTCCT 1380  
 DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCGCTGTCTTCTTGAAGCATTCCT 1380  
 QY 1381 AGGGGTCTTCCCTCTCGCCAAAGGAATGAAAGTCTGTTGAATGCTGGAAGAGCA 1440  
 DB 1381 AGGGGTCTTCCCTCTCGCCAAAGGAATGAAAGTCTGTTGAATGCTGGAAGAGCA 1440  
 QY 1441 GTTCCTCTGAAGAGTTCTTGAAGAGCAACACGCTGTAGGAGCCCTTTGACGAGCGG 1500  
 DB 1441 GTTCCTCTGAAGAGTTCTTGAAGAGCAACACGCTGTAGGAGCCCTTTGACGAGCGG 1500

QY 1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCCAAAGCCAGTGTATAGATACACCT 1560  
DB 1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCCAAAGCCAGTGTATAGATACACCT 1560  
QY 1561 GCAAGGCGGCAACAACCCAGTGCAGTGTGTGAGTTGGATAGTGTGGAAAGAGTCAAA 1620  
DB 1561 GCAAGGCGGCAACAACCCAGTGCAGTGTGTGAGTTGGATAGTGTGGAAAGAGTCAAA 1620  
QY 1621 TGGCTCTCCTCAAGCGTATTCAACAAGGGCTGAAGTATGCCAAGAGTACCCCATGT 1680  
DB 1621 TGGCTCTCCTCAAGCGTATTCAACAAGGGCTGAAGTATGCCAAGAGTACCCCATGT 1680  
QY 1681 ATGGGATCTCATCTGGGCTCTGGTGCACATGCTTTTACATGTGTAGTGCAGGTAAAA 1740  
DB 1681 ATGGGATCTCATCTGGGCTCTGGTGCACATGCTTTTACATGTGTAGTGCAGGTAAAA 1740  
QY 1741 AAGCTCTAGGCCCCCGAACAACGGGAGCTGGTTTTTCCTTTGAATAACAGTAAATACC 1800  
DB 1741 AAGCTCTAGGCCCCCGAACAACGGGAGCTGGTTTTTCCTTTGAATAACAGTAAATACC 1800  
QY 1801 ATGGGCTATTAGGCTTACTCCCAACAGAGGGAGGCTACTTGGCTGCATCATCACT 1860  
DB 1801 ATGGGCTATTAGGCTTACTCCCAACAGAGGGAGGCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCTCAACAGGCGGACAGAAACAGTGCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
DB 1861 AGCTCAACAGGCGGACAGAAACAGTGCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACAAATCTTCTGGGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTC 1980  
DB 1921 ACACAAATCTTCTGGGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTC 1980  
QY 1981 GGTCTAAAGACCTTGGCGGCCCAAGGGCCCAATCACCAAAATGTACACCAATGTGGAC 2040  
DB 1981 GGTCTAAAGACCTTGGCGGCCCAAGGGCCCAATCACCAAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACCATGCACTGC 2100  
DB 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACCATGCACTGC 2100  
QY 2101 GGCAGCTCGGACCTTACTTGGTCAAGGACGATGCGATGTCATTCGGGTGCGCGCGG 2160  
DB 2101 GGCAGCTCGGACCTTACTTGGTCAAGGACGATGCGATGTCATTCGGGTGCGCGCGG 2160  
QY 2161 GCGCAGCAGGCGGAGCTACTCTCCCGAGCCCGTCTCTACTTGAAGGCTCTTCG 2220  
DB 2161 GCGCAGCAGGCGGAGCTACTCTCCCGAGCCCGTCTCTACTTGAAGGCTCTTCG 2220  
QY 2221 GCGGTCCTACTGCTCTGCCCTCGGGCAGCTGTGGGATCTTTCGGGCTGCGGTGTC 2280  
DB 2221 GCGGTCCTACTGCTCTGCCCTCGGGCAGCTGTGGGATCTTTCGGGCTGCGGTGTC 2280  
QY 2281 ACCGAGGGGTTGCAAGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACCACTATG 2340  
DB 2281 ACCGAGGGGTTGCAAGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCAGGCAACTGTCCTCCCGAGCCCGTCTCTACTTGAAGGCTCTTCG 2400  
DB 2341 CGGTCCCGGCTTTCAGGCAACTGTCCTCCCGAGCCCGTCTCTACTTGAAGGCTCTTCG 2400  
QY 2401 GCGCATCTACGCGCCCTACTGCTAGCGGCAAGACACTAAGTGTGCGGCTGCGGTATGCA 2460  
DB 2401 GCGCATCTACGCGCCCTACTGCTAGCGGCAAGACACTAAGTGTGCGGCTGCGGTATGCA 2460  
QY 2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCCCTAGGTTTCGGG 2520  
DB 2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCCCTAGGTTTCGGG 2520  
QY 2521 GGGTATATGCTTAAGGCACATGGTATCGACCCCTAACATCAGAACCGGGGTAAAGCCATC 2580  
DB 2521 GGGTATATGCTTAAGGCACATGGTATCGACCCCTAACATCAGAACCGGGGTAAAGCCATC 2580  
QY 2581 ACCACGGGTGCCCCATCAGTACTCCACCTATGCAAGTCTTCTGCGGACGGTGTGC 2640

DB 2581 ACCACGGGTGCCCCATCAGTACTCCACCTATGCAAGTCTTCTGCGGACGGTGTGC 2640  
QY 2641 TCTGGGGCGCCCTATGACATCATATAATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB 2641 TCTGGGGCGCCCTATGACATCATATAATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCAATCGGCAAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGGCTCGTGGT 2760  
DB 2701 ATCTGGGCAATCGGCAAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGGCTCGTGGT 2760  
QY 2761 CTGGCCACCGCTAGCGCTCGGGATCGGTTCACCGTGCACATCCAAATATCGAGAGGT 2820  
DB 2761 CTGGCCACCGCTAGCGCTCGGGATCGGTTCACCGTGCACATCCAAATATCGAGAGGT 2820  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
DB 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGCACTCTCAATTTCTGCCAATCCAAAGAGAAATGTGATGAGCTCGCGCG 2940  
DB 2881 AAGGGGGGAGGCACTCTCAATTTCTGCCAATCCAAAGAGAAATGTGATGAGCTCGCGCG 2940  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATATCGTGC 3000  
DB 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATATCGTGC 3000  
QY 3001 ATACCAACTAGCGGACAGCTCATTGTCTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
DB 3001 ATACCAACTAGCGGACAGCTCATTGTCTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
QY 3061 GGCATTTCCAGTCACTGATGCACTGCAATATCATGTGTCCACGACAGTCTGACTTCAGC 3120  
DB 3061 GGCATTTCCAGTCACTGATGCACTGCAATATCATGTGTCCACGACAGTCTGACTTCAGC 3120  
QY 3121 CTGGACCGGACCTTACCAATTTGAGAGCAGCGGTGCCAAGACCGGGTGTCAACGCTG 3180  
DB 3121 CTGGACCGGACCTTACCAATTTGAGAGCAGCGGTGCCAAGACCGGGTGTCAACGCTG 3180  
QY 3181 CAGCGCGAGGAGGACTGCTAGGGGACAGTGGGCATTTACAGGTTTCTGACTCCAGCA 3240  
DB 3181 CAGCGCGAGGAGGACTGCTAGGGGACAGTGGGCATTTACAGGTTTCTGACTCCAGCA 3240  
QY 3241 GAAACGCGCTCGGGCATGTTTCGATTTCTGCGAGTGTCTGAGTGTATACGCGGCTGT 3300  
DB 3241 GAAACGCGCTCGGGCATGTTTCGATTTCTGCGAGTGTCTGAGTGTATACGCGGCTGT 3300  
QY 3301 GCTTGGTAGAGCTCAGCGCCCGGAGACCTCAGTTAGTGGTGGGCTTACCTAAACACA 3360  
DB 3301 GCTTGGTAGAGCTCAGCGCCCGGAGACCTCAGTTAGTGGTGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGGTTGCCGCTGCGCAGGACCATCTGGAGTCTGGGAGAGGCTTTTACAGGCTC 3420  
DB 3361 CCAGGGTTGCCGCTGCGCAGGACCATCTGGAGTCTGGGAGAGGCTTTTACAGGCTC 3420  
QY 3421 ACCCAATAGACGCCCATTTCTTGTCCAGACTAAGACGCGAGGAGCAACTTCCCTTAC 3480  
DB 3421 ACCCAATAGACGCCCATTTCTTGTCCAGACTAAGACGCGAGGAGCAACTTCCCTTAC 3480  
QY 3481 CTGTGTAGTACACAGCTCAGGTGTGCGCAGGGCTCAGGTCTACCTCATCTGTGGAC 3540  
DB 3481 CTGTGTAGTACACAGCTCAGGTGTGCGCAGGGCTCAGGTCTACCTCATCTGTGGAC 3540  
QY 3541 CAAATGTGGAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGGCTTCTTACAGGCTC 3600  
DB 3541 CAAATGTGGAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGGCTTCTTACAGGCTC 3600  
QY 3601 TATAGGCTGGGAGCGGTTCAAAACGAGGTACTACACACACCCCATAAACCAATATCATC 3660  
DB 3601 TATAGGCTGGGAGCGGTTCAAAACGAGGTACTACACACACCCCATAAACCAATATCATC 3660  
QY 3661 ATGGCATCATGTGGGCTGAGCTGAGGTGTGACGAGCACTGGGTGTGTAGGCGGA 3720

Db 3661 ATGSCATGCTCGGCTGACCTGAGTGCTGACGAGCACCTGGTGCTGGTAGGGGA 3720  
Qy 3721 GTCCTAGAGCTCTGGCGCGGTATTGCTGTGACAAAGCAGCGTGGTCAITTTGTGGCAGG 3780  
Db 3721 GTCCTAGAGCTCTGGCGCGGTATTGCTGTGACAAAGCAGCGTGGTCAITTTGTGGCAGG 3780  
Qy 3781 ATCATCTGTCCGGAAGCCGCGCATCATCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTGTCCGGAAGCCGCGCATCATCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Qy 3841 GATGAGTGAAGAGTGGCTCTCACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
Db 3841 GATGAGTGAAGAGTGGCTCTCACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
Qy 3901 GAACAATTCAAACAGAGGGAATCGGTTGCTGMAACAGCCACCAAGCAGCGGAGCT 3960  
Db 3901 GAACAATTCAAACAGAGGGAATCGGTTGCTGMAACAGCCACCAAGCAGCGGAGCT 3960  
Qy 3961 GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGCTTCTGGCGGAAGCATATG 4020  
Db 3961 GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGCTTCTGGCGGAAGCATATG 4020  
Qy 4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Qy 4081 GCGATAGCATCACTGATGGCATCAAGCTCTATCAGCAGCCGCTCACCACCCAAAT 4140  
Db 4081 GCGATAGCATCACTGATGGCATCAAGCTCTATCAGCAGCCGCTCACCACCCAAAT 4140  
Qy 4141 ACCCTCTGTTTAAATCTGGGGGATGGGTGGCGGCCCACTTGTCTCCCAAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAATCTGGGGGATGGGTGGCGGCCCACTTGTCTCCCAAGCGCT 4200  
Qy 4201 GCTTCTGCTTTCGTAGGCGCCGCTATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCCGCTATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGG 4260  
Qy 4261 AAGGTGCTTGGATATTTTGGCAGGTTATGAGCAGGCGTGGCAGGCGCTCGTGCC 4320  
Db 4261 AAGGTGCTTGGATATTTTGGCAGGTTATGAGCAGGCGTGGCAGGCGCTCGTGCC 4320  
Qy 4321 TTTAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT 4380  
Db 4321 TTTAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT 4380  
Qy 4381 ATCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTGCGAGCGCATCTGCTGCGCAC 4440  
Db 4381 ATCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTGCGAGCGCATCTGCTGCGCAC 4440  
Qy 4441 GTGGGCCGAGGGGAGGGGCTGTGAGTGAAGACCGGCTGATAGCGTTGCTTTCGGG 4500  
Db 4441 GTGGGCCGAGGGGAGGGGCTGTGAGTGAAGACCGGCTGATAGCGTTGCTTTCGGG 4500  
Qy 4501 GGTAAACAGCTCTCCCGCCAGCCTATGTGCTGAGAGGAGCGCTGCAGCAGTGTCACT 4560  
Db 4501 GGTAAACAGCTCTCCCGCCAGCCTATGTGCTGAGAGGAGCGCTGCAGCAGTGTCACT 4560  
Qy 4561 CAGATCTCTAGTCTTACCATCACTAGCTGTGAAGAGGTTCCAGTGGATCAAC 4620  
Db 4561 CAGATCTCTAGTCTTACCATCACTAGCTGTGAAGAGGTTCCAGTGGATCAAC 4620  
Qy 4621 GAGGAGTCTCCAGCGCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Db 4621 GAGGAGTCTCCAGCGCATGCTCCGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Qy 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATTCGCGGA 4740  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATTCGCGGA 4740  
Qy 4741 GTCCCTCTCTCTCATGTCAAGTGGGTACAGGAGTCTGGCGGGGACCGCATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCAAGTGGGTACAGGAGTCTGGCGGGGACCGCATCATG 4800

Qy 4801 CAAACACCTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
Db 4801 CAAACACCTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
Qy 4861 ATCGTGGGGCTAGGACCTGTAGTAAACGCTGGGATGGAAATTTCCCATTAACGCTAC 4920  
Db 4861 ATCGTGGGGCTAGGACCTGTAGTAAACGCTGGGATGGAAATTTCCCATTAACGCTAC 4920  
Qy 4921 ACCAGGGCCCTCGACGCTCCCGGCGCAAAATTTCTAGGGCTGTGGGGGTG 4980  
Db 4921 ACCAGGGCCCTCGACGCTCCCGGCGCAAAATTTCTAGGGCTGTGGGGGTG 4980  
Qy 4981 GCTGCTGAGGAGTACGTGGAGTTCACGCGGTGGGGATTTCCCATGACGCGCATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGTTCACGCGGTGGGGATTTCCCATGACGCGCATG 5040  
Qy 5041 ACCACTGAACAGTAAAGTGCCTGTCAGGTTCCGGCCCGCAATTTCTTACAGAGTG 5100  
Db 5041 ACCACTGAACAGTAAAGTGCCTGTCAGGTTCCGGCCCGCAATTTCTTACAGAGTG 5100  
Qy 5101 GATGGGTGCGGTGTCACAGGTACGCTCCAGCTGCAAAACCTCTCTACGGGAGGAGTGC 5160  
Db 5101 GATGGGTGCGGTGTCACAGGTACGCTCCAGCTGCAAAACCTCTCTACGGGAGGAGTGC 5160  
Qy 5161 ACATTTCTGTCGGGCTCAATCAATACCTGTTGGGTCAACAGTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTTCTGTCGGGCTCAATCAATACCTGTTGGGTCAACAGTCCCATGCGAGCCGAA 5220  
Qy 5221 CCGAGCTAGCAGTGTCTCACTTCCATGCTACGAGTCCGAGTCCCATGCGGAGGAG 5280  
Db 5221 CCGAGCTAGCAGTGTCTCACTTCCATGCTACGAGTCCGAGTCCCATGCGGAGGAG 5280  
Qy 5281 GCTAAGCTTAGCTGGCCAGGAGTCTCCCGCTCTTGGCGAGTCACTACGTAGCCAG 5340  
Db 5281 GCTAAGCTTAGCTGGCCAGGAGTCTCCCGCTCTTGGCGAGTCACTACGTAGCCAG 5340  
Qy 5341 CTGCTGCGCTTCTTGAAGCAACATGCTACCTCGTCACTACCTCCCGGACGCTGAC 5400  
Db 5341 CTGCTGCGCTTCTTGAAGCAACATGCTACCTCGTCACTACCTCCCGGACGCTGAC 5400  
Qy 5401 CTCATCGAGGCAACCTCTGTGGCGGAGAGATGGCGGGGAAATCACCCCGTGGAG 5460  
Db 5401 CTCATCGAGGCAACCTCTGTGGCGGAGAGATGGCGGGGAAATCACCCCGTGGAG 5460  
Qy 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGGAGGATGAG 5520  
Db 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGGAGGATGAG 5520  
Qy 5521 AGGGAAGTATCCGTTCCCGGAGATCTTGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCCGGAGATCTTGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
Qy 5581 CCCATATGGGCAACGCGCGGATTAACACCTTCCAGTCTTTCGAGCGCTCCAGGCGGAC 5640  
Db 5581 CCCATATGGGCAACGCGCGGATTAACACCTTCCAGTCTTTCGAGCGCTCCAGGCGGAC 5640  
Qy 5641 TAGCTCCCTCCAGTGTACAGGGTGTCCATTCGCGCTCCAAAGCGGCTCCGATACCA 5700  
Db 5641 TAGCTCCCTCCAGTGTACAGGGTGTCCATTCGCGCTCCAAAGCGGCTCCGATACCA 5700  
Qy 5701 CTTCCACGAGGAGGAGCGGTGTCTGTGTCAGAAATTCAGTGTCTTCTGCTTGGG 5760  
Db 5701 CTTCCACGAGGAGGAGCGGTGTCTGTGTCAGAAATTCAGTGTCTTCTGCTTGGG 5760  
Qy 5761 GAGCTCGCACAAAGACCTTCGCGAGTCCGAATTCGTCCGCTCCAGCGGACCGCA 5820  
Db 5761 GAGCTCGCACAAAGACCTTCGCGAGTCCGAATTCGTCCGCTCCAGCGGACCGCA 5820  
Qy 5821 ACGGCTCTCTGACCGAGCCCTCCGACGCGGAGTCCGAGTCCGAGTTCGATCGTAC 5880  
Db 5821 ACGGCTCTCTGACCGAGCCCTCCGACGCGGAGTCCGAGTTCGAGTTCGATCGTAC 5880

Qy	5881	TCCTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGTCTTGG	5941
Db	5881	TCCTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGTCTTGG	5940
Qy	5941	TCACCCGTAGCGAGGAGCTAGTGAGGAGCTGCTCTGCTGCTCGATGTCCTACACATGG	6000
Db	5941	TCACCCGTAGCGAGGAGCTAGTGAGGAGCTGCTCTGCTGCTCGATGTCCTACACATGG	6000
Qy	6001	ACAGCGCCCTTGATCAGCCATCGCGTTCGGAGGAAAACCAAGCTGCCCATCAATGCACATG	6060
Db	6001	ACAGCGCCCTTGATCAGCCATCGCGTTCGGAGGAAAACCAAGCTGCCCATCAATGCACATG	6060
Qy	6061	AGAACTCTTTGTCCTGCTACCAAACTTTGGTCTATGCTACAACTCTCGCAGCGCAAGC	6120
Db	6061	AGAACTCTTTGTCCTGCTACCAAACTTTGGTCTATGCTACAACTCTCGCAGCGCAAGC	6120
Qy	6121	CTGGCGCAGAGAGAGTCACTTTGACAGACTCGAGTCTCTGGACGACCACTACCGGGAC	6180
Db	6121	CTGGCGCAGAGAGAGTCACTTTGACAGACTCGAGTCTCTGGACGACCACTACCGGGAC	6180
Qy	6181	GTGCTCAAGGAGATGAAGGGGAAAGCGCTGCACAGCTTAAAGCTTAAAGCTTAACTTCGTGGAG	6240
Db	6181	GTGCTCAAGGAGATGAAGGGGAAAGCGCTGCACAGCTTAAAGCTTAAAGCTTAACTTCGTGGAG	6240
Qy	6241	GAGCCTGTGAAGTGAAGCGCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Db	6241	GAGCCTGTGAAGTGAAGCGCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Qy	6301	GAGTCCGGAACTATCCAGCAGCGCGTTAAACACATCCGCTCCGTGTGGAGGACTTG	6360
Db	6301	GAGTCCGGAACTATCCAGCAGCGCGTTAAACACATCCGCTCCGTGTGGAGGACTTG	6360
Qy	6361	CTGGAAGACACTGAGACACCAATTGACACCAACATCATGCAAAAAATGAGTTTCTGC	6420
Db	6361	CTGGAAGACACTGAGACACCAATTGACACCAACATCATGCAAAAAATGAGTTTCTGC	6420
Qy	6421	GTCCAAACGAGAGAGGGGGCGGAAAGCCAGCTGCGCTTATCGTATTCACAGATTTGGGG	6480
Db	6421	GTCCAAACGAGAGAGGGGGCGGAAAGCCAGCTGCGCTTATCGTATTCACAGATTTGGGG	6480
Qy	6481	GTTCGTGTGGCAGAAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG	6540
Db	6481	GTTCGTGTGGCAGAAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG	6540
Qy	6541	ATGGGCTCTTCAACGGAATCCCAATCTCTCCCTGGACAGCGGTTCGAGTTCCTGTGTAAT	6600
Db	6541	ATGGGCTCTTCAACGGAATCCCAATCTCTCCCTGGACAGCGGTTCGAGTTCCTGTGTAAT	6600
Qy	6601	GCCTGGAAACGGAATAAGCCCTATGGGCTTCGCATATGACACCGCGTGTTTGACTCA	6660
Db	6601	GCCTGGAAACGGAATAAGCCCTATGGGCTTCGCATATGACACCGCGTGTTTGACTCA	6660
Qy	6661	ACGCTCACTCAGAAATGACATCCGCTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC	6720
Db	6661	ACGCTCACTCAGAAATGACATCCGCTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC	6720
Qy	6721	CCGGAAGCCAGACAGGCCATAAGGTCCCTCAGAGGCGCTTTACATCGGGGCCCCCTTG	6780
Db	6721	CCGGAAGCCAGACAGGCCATAAGGTCCCTCAGAGGCGCTTTACATCGGGGCCCCCTTG	6780
Qy	6781	ACTAAATCTTAAAGGCGAAGACTCGCGGTATCGCCGCTATCGCGCGCGCGTGTGACTGACG	6840
Db	6781	ACTAAATCTTAAAGGCGAAGACTCGCGGTATCGCCGCTATCGCGCGCGCGTGTGACTGACG	6840
Qy	6841	ACCAGCTCGCGTAATACCTTCAATGTACTTGAAGCGCGCTCGGCGCTGTGAGCTGGG	6900
Db	6841	ACCAGCTCGCGTAATACCTTCAATGTACTTGAAGCGCGCTCGGCGCTGTGAGCTGGG	6900
Qy	6901	AAGCTCCAGGACTGCACATGCTCGTATGCGGAGACGACCTTCTCGTTACTGTGAAGC	6960
Db	6901	AAGCTCCAGGACTGCACATGCTCGTATGCGGAGACGACCTTCTCGTTACTGTGAAGC	6960
Qy	6961	GCGGGGACCAAGAGGACGAGCGAGCCTTACGGGCCCTTACGGAGGCTATGACTAGATAC	7020

[illegible]



AAD25326  
ID AAD25326 standard; cDNA; 7989 BP.  
XX AC AAD25326;  
XX 12-MAR-2002 (first entry)  
XX DE Hepatitis C virus (HCV) adaptive replicon V mutant cDNA.  
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KW adaptive replicon V; mutant; ss.  
XX Hepatitis C virus.  
OS Synthetic.  
XX  
FH Location/Qualifiers  
FT 4642..5982  
FT /\*tag= a  
FT /product= "NSSA variant of HCV adaptive replicon v"  
FT /note= "CDS does not include both start and stop codon"  
FT /partial  
FT mutation  
FT replace(5320, G)  
FT /\*tag= b  
XX  
PN WO200189364-A2.  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-US016822.  
XX  
XX 23-MAY-2000; 2000US-00576989.  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Rice CM, Blight KJ;  
XX  
XX WPI; 2002-066755/09.  
XX P-PSDB; AAE15724.  
XX  
XX Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
XX Disclosure; Page 80-83; 174pp; English.  
XX  
XX The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
XX The present sequence is Hepatitis C virus (HCV) adaptive replicon V

CC mutant cDNA. This sequence is generated by the mutation g to t at  
CC position 5320 of HCVreplbBartMan/Avail cDNA  
XX  
XX Sequence 7989 BP; 1647 A; 2369 C; 2241 G; 1732 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 7987.4; DB 6; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCCAGCCCCCGAATTGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60  
DB 1 GCCAGCCCCCGAATTGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60  
QY 61 TCTTCAGCGCAAGAGCGTCTAGCCATGCGCTTAGTATGAGTGTCTGTCAGCTCCAGGAC 120  
DB 61 TCTTCAGCGCAAGAGCGTCTAGCCATGCGCTTAGTATGAGTGTCTGTCAGCTCCAGGAC 120  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCCAG 180  
DB 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCCAG 180  
QY 181 GACGACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTCCGCC 240  
DB 181 GACGACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTCCGCC 240  
QY 241 CGGAGACTGCTAGCGGAGTGTGTTGGTCTGCGAAAGGCTTGTGGTACTGCTGATAGG 300  
DB 241 CGGAGACTGCTAGCGGAGTGTGTTGGTCTGCGAAAGGCTTGTGGTACTGCTGATAGG 300  
QY 301 GTGCTTGGAGTGGCCCGGAGGTTCTGTAGCCGTCACCATGACGACGATCCCTAAAC 360  
DB 301 GTGCTTGGAGTGGCCCGGAGGTTCTGTAGCCGTCACCATGACGACGATCCCTAAAC 360  
QY 361 CTCGAAGAAACCAAGGCGCGCATGATGACCAAGATGATGACGACGAGTTCTC 420  
DB 361 CTCGAAGAAACCAAGGCGCGCATGATGACCAAGATGATGACGACGAGTTCTC 420  
QY 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGACTGGGCGACAAACAGACAATCGCTGT 480  
DB 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGACTGGGCGACAAACAGACAATCGCTGT 480  
QY 481 CTGATGCGCGGTGTTCGGTGTGACGCGAGGGCGCGCGGTTCTTTTGTCAAGACG 540  
DB 481 CTGATGCGCGGTGTTCGGTGTGACGCGAGGGCGCGCGGTTCTTTTGTCAAGACG 540  
QY 541 ACCTGTCGGTGCCTGTAATGAACTGACGAGGAGCGGCGGCTATCGTGGCTGGCCA 600  
DB 541 ACCTGTCGGTGCCTGTAATGAACTGACGAGGAGCGGCGGCTATCGTGGCTGGCCA 600  
QY 601 CGACGGCGGTTCTTTCGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
DB 601 CGACGGCGGTTCTTTCGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
QY 661 TGCTATTGGGGAAGTGTGCGGGGAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGGAGA 720  
DB 661 TGCTATTGGGGAAGTGTGCGGGGAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGGAGA 720  
QY 721 AAGTATCATCATGCTGATGCAATGCGCGGCTGATACGCTTGTATCGGCTACTGCG 780  
DB 721 AAGTATCATCATGCTGATGCAATGCGCGGCTGATACGCTTGTATCGGCTACTGCG 780  
QY 781 CATTCGACCAACCAAGCGAATCATGCTGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
DB 781 CATTCGACCAACCAAGCGAATCATGCTGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
QY 841 TTGTGATCAGATGATGTGACGAGAGCATCAGGGGTCCGCGCAGCGAATCTTTCG 900  
DB 841 TTGTGATCAGATGATGTGACGAGAGCATCAGGGGTCCGCGCAGCGAATCTTTCG 900  
QY 901 CCAGGCTCAAGGCGGCGATGCGCGGAGGATCTCGTCTGACCCATGCGGATGCGCT 960  
DB 901 CCAGGCTCAAGGCGGCGATGCGCGGAGGATCTCGTCTGACCCATGCGGATGCGCT 960



Qy 961 GCTTCCGAATATCATGTTGAAATGCGCTTTCTGGATTCATCGACTGTGCGCGC 1020  
Db 961 GCTTGGCGAATATCATGTTGAAATGCGCGTTTCTGGATTCATCGACTGTGCGCGC 1020  
Qy 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080  
Db 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080  
Qy 1081 TTGCGCGGAATGGCTGACCGCTTCCTCGTCTTACGGTATCGCGCTCCGATTCGC 1140  
Db 1081 TTGCGCGGAATGGCTGACCGCTTCCTCGTCTTACGGTATCGCGCTCCGATTCGC 1140  
Qy 1141 AGCGATCGCCTTCTATCGCCTTCTGACGAGTTCTTCTGAGTTTAAACAGACACACAGC 1200  
Db 1141 AGCGATCGCCTTCTATCGCCTTCTGACGAGTTCTTCTGAGTTTAAACAGACACACAGC 1200  
Qy 1201 GTTTCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCCGCCCTTAAGCTTACTGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCCGCCCTTAAGCTTACTGGC 1260  
Qy 1261 CGAAGCGCTTGAATAGCGCGTGTGCTTCTATATGTTTATGTTTCCACCATATG 1320  
Db 1261 CGAAGCGCTTGAATAGCGCGTGTGCTTCTATATGTTTATGTTTCCACCATATG 1320  
Qy 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCTGTCTTCTTGAAGAGCATTCCT 1380  
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCTGTCTTCTTGAAGAGCATTCCT 1380  
Qy 1381 AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGCTGTGAAGGAAGCA 1440  
Db 1381 AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGCTGTGAAGGAAGCA 1440  
Qy 1441 GTTCTCTGGAAGTCTTCTGAAGCAACCAACGTCTGTAGGACCTTTGAGGACGCGG 1500  
Db 1441 GTTCTCTGGAAGTCTTCTGAAGCAACCAACGTCTGTAGGACCTTTGAGGACGCGG 1500  
Qy 1501 AACCCGCCACTGTGCGACAGGTGCTCTGCGGCCAAAGACACGTGATAGATACACCT 1560  
Db 1501 AACCCGCCACTGTGCGACAGGTGCTCTGCGGCCAAAGACACGTGATAGATACACCT 1560  
Qy 1561 GAAAGCGGCAAAACCCAGTCCAGTGTGAGTTGGATGTTGTGAAAGAGTCAAA 1620  
Db 1561 GAAAGCGGCAAAACCCAGTCCAGTGTGAGTTGGATGTTGTGAAAGAGTCAAA 1620  
Qy 1621 TGCTCTCTCAAGCTATTCACAGGCGCTGAAGATCCCAAGAGTACCCATGT 1680  
Db 1621 TGCTCTCTCAAGCTATTCACAGGCGCTGAAGATCCCAAGAGTACCCATGT 1680  
Qy 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAA 1740  
Qy 1741 AAGCTTAGGCCCCCGAACCAAGGACCGTGTGTTTCTTTGAAACACGATAATACC 1800  
Db 1741 AAGCTTAGGCCCCCGAACCAAGGACCGTGTGTTTCTTTGAAACACGATAATACC 1800  
Qy 1801 ATGGCGCTATTACGCGCTACTCCACAGACCGGAGGCTACTTGGTGATCATCACT 1860  
Db 1801 ATGGCGCTATTACGCGCTACTCCACAGACCGGAGGCTACTTGGTGATCATCACT 1860  
Qy 1861 AGCCTCACAGCGGGACAGGAACAGGTGCGAGGGGAGGTCCAGTGGTCTCCACCGCA 1920  
Db 1861 AGCCTCACAGCGGGACAGGAACAGGTGCGAGGGGAGGTCCAGTGGTCTCCACCGCA 1920  
Qy 1921 ACACAATCTTCTGCGGACCTCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
Db 1921 ACACAATCTTCTGCGGACCTCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
Qy 1981 GGTCAAGACCTTCCCGGCCAAAGGCGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 1981 GGTCAAGACCTTCCCGGCCAAAGGCGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Qy 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGGCTTCTTGTGACACCATGCACCTGC 2100

Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGGCTTCTTGTGACACCATGCACCTGC 2100  
Qy 2101 GCGAGCTCGGACCTTTACTTGGTCAAGGAGCATCGGATGTCATTCGGGTGCGCGCGG 2160  
Db 2101 GCGAGCTCGGACCTTTACTTGGTCAAGGAGCATCGGATGTCATTCGGGTGCGCGCGG 2160  
Qy 2161 GCGCAACAGAGGGGAGCCTACTCTCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GCGCAACAGAGGGGAGCCTACTCTCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Qy 2221 GCGGCTCCACTGCTCTCGCCCTCGGSCACGCTGTGGGCATCTTTTGGGCTGCGGTGTC 2280  
Db 2221 GCGGCTCCACTGCTCTCGCCCTCGGSCACGCTGTGGGCATCTTTTGGGCTGCGGTGTC 2280  
Qy 2281 ACCCGAGGGGTGCGAAGGGGCTGACTTTGTACCCGTCGAGTCTATGGAAACCACTATG 2340  
Db 2281 ACCCGAGGGGTGCGAAGGGGCTGACTTTGTACCCGTCGAGTCTATGGAAACCACTATG 2340  
Qy 2341 CGGTCCCGGCTTTCACGGAACAACCTCGTCCCTCCGCCGCTACCGCAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCACGGAACAACCTCGTCCCTCCGCCGCTACCGCAGACATTCAGGTG 2400  
Qy 2401 GCGCATCTACAGCCCTTACTGTAGCGGCAAGAGCATTAAGGTGCGGCTGCGGTATGCA 2460  
Db 2401 GCGCATCTACAGCCCTTACTGTAGCGGCAAGAGCATTAAGGTGCGGCTGCGGTATGCA 2460  
Qy 2461 GCCCAAGGGTATAGGTGCTTGTCTGAACCCGTCGCGCCACCCCTAGTTTCGGG 2520  
Db 2461 GCCCAAGGGTATAGGTGCTTGTCTGAACCCGTCGCGCCACCCCTAGTTTCGGG 2520  
Qy 2521 GCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTAAAGACCATC 2580  
Db 2521 GCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTAAAGACCATC 2580  
Qy 2581 ACCAGCGGTGCGCCCATCAGTACTCACCTATGGCAAGTTTCTGCCGAGGTGGTGTG 2640  
Db 2581 ACCAGCGGTGCGCCCATCAGTACTCACCTATGGCAAGTTTCTGCCGAGGTGGTGTG 2640  
Qy 2641 TCTGGGGCGCTTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCTTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Qy 2701 ATCTGGGCATCGGCACAGTCTTGGACCAAGCGGAGACGCTGAGCGGACCTCGTGTG 2760  
Db 2701 ATCTGGGCATCGGCACAGTCTTGGACCAAGCGGAGACGCTGAGCGGACCTCGTGTG 2760  
Qy 2761 CTGCGCACCGCTACGCTCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG 2820  
Db 2761 CTGCGCACCGCTACGCTCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG 2820  
Qy 2821 GCTCTCTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTCTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Qy 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCAGAGAAATGTGATGAGTCTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCAGAGAAATGTGATGAGTCTCGCGCG 2940  
Qy 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCCGTC 3000  
Qy 3001 ATACCAACTAGCGGAGACGTCAATGCTGTAGCAACGAGCGCTCTAATGACGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGGAGACGTCAATGCTGTAGCAACGAGCGCTCTAATGACGGCTTTACC 3060  
Qy 3061 GCGATTTTCAGTCACTGATCGACTGCAATACATGTGTCCAGACAGTTCAGCTTCAGC 3120  
Db 3061 GCGATTTTCAGTCACTGATCGACTGCAATACATGTGTCCAGACAGTTCAGCTTCAGC 3120  
Qy 3121 CTGGACCGGACCTTCAACATTCAGAGGACGACCGGTGCCAAGACGCGGTGTACGCTCG 3180

Db 3121 CTGGACCGACCTTACACATTGAGACGACGACCGGTGCCACAGACGCGGTGTACGCTCG 3180  
Qy 3181 CAGCGCGAGGACGAGACTGGTAGGGGACGAGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGCGAGGACGAGACTGGTAGGGGACGAGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
Qy 3241 GAAACGCGCTCGGCACTGTTCGATTCCTCGGTTCTGTCGAGTGTATGACGCGGGCTGT 3300  
Db 3241 GAAACGCGCTCGGCACTGTTCGATTCCTCGGTTCTGTCGAGTGTATGACGCGGGCTGT 3300  
Qy 3301 GCTTGGTACGAGCTCACGCGCCCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGGTACGAGCTCACGCGCCCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
Qy 3361 CCAGGGTTGCCGTCTGCGAGAACCATCTGAGAGTTCTGGGAGAGGTTCTTTACAGGCTC 3420  
Db 3361 CCAGGGTTGCCGTCTGCGAGAACCATCTGAGAGTTCTGGGAGAGGTTCTTTACAGGCTC 3420  
Qy 3421 ACCCACATAGACGCCATTTCTGTCCAGACTAAGACGAGGAGACAACTTCCCTAC 3480  
Db 3421 ACCCACATAGACGCCATTTCTGTCCAGACTAAGACGAGGAGACAACTTCCCTAC 3480  
Qy 3481 CTGCTAGCATACAGGCTACGGTGTGCGCGAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGCTAGCATACAGGCTACGGTGTGCGCGAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Qy 3541 CAAATGTGAAATGTCTCATAGGCTAAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCT 3600  
Db 3541 CAAATGTGAAATGTCTCATAGGCTAAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCT 3600  
Qy 3601 TATAGCTGGGAGCGCTTCAAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
Db 3601 TATAGCTGGGAGCGCTTCAAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
Qy 3661 ATGGCATGATGTCGCTGACCTGGAGTGTGTCACGAGACCTGGGTGTGTAGGCGGA 3720  
Db 3661 ATGGCATGATGTCGCTGACCTGGAGTGTGTCACGAGACCTGGGTGTGTAGGCGGA 3720  
Qy 3721 GTCTAGACGCTTGGCGGGTATTCGCTGAACAGGAGAGTGTGTATTTGGGACAG 3780  
Db 3721 GTCTAGACGCTTGGCGGGTATTCGCTGAACAGGAGAGTGTGTATTTGGGACAG 3780  
Qy 3781 ATCATCTTGTCCGGAAGCGGCCATTCATCCCGACAGGGAATTCCTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGGAAGCGGCCATTCATCCCGACAGGGAATTCCTTACCGGAGTTC 3840  
Qy 3841 GATGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCAAGCTCGCC 3900  
Db 3841 GATGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCAAGCTCGCC 3900  
Qy 3901 GAACAAATTCAGAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT 3960  
Db 3901 GAACAAATTCAGAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT 3960  
Qy 3961 GCTGCTCCGCTGGTGAATTCAGAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGAATTCAGAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Qy 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
Qy 4081 GCGATAGCATCATGATGCAATTCAGGCTCTTACACAGCGGCTCACACCCACAT 4140  
Db 4081 GCGATAGCATCATGATGCAATTCAGGCTCTTACACAGCGGCTCACACCCACAT 4140  
Qy 4141 ACCCTCTGTTTAACTCTCGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAACTCTCGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
Qy 4201 GCTTCTGCTTTCGTAGCGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGCGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260

Qy 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGCGCTCGTGGCC 4320  
Db 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGCGCTCGTGGCC 4320  
Qy 4321 TTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGT 4380  
Db 4321 TTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGT 4380  
Qy 4381 ATCCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTCGCGCAGGATACTGCTGGCAC 4440  
Db 4381 ATCCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTCGCGCAGGATACTGCTGGCAC 4440  
Qy 4441 GTGGGCCACGAGGAGGCGCTGTGCACTGGATGAACCGGCTGATAGCGTTCCGCGG 4500  
Db 4441 GTGGGCCACGAGGAGGCGCTGTGCACTGGATGAACCGGCTGATAGCGTTCCGCGG 4500  
Qy 4501 GGTAAACCACTCTCCCGCAGCTATGTGCTGAGAGCGCTGCGACGCTGTCACT 4560  
Db 4501 GGTAAACCACTCTCCCGCAGCTATGTGCTGAGAGCGCTGCGACGCTGTCACT 4560  
Qy 4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
Qy 4621 GAGGACTGTCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATGATGC 4680  
Db 4621 GAGGACTGTCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATGATGC 4680  
Qy 4681 ACGGTTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTGCGCGATTTCCGCGGA 4740  
Db 4681 ACGGTTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTGCGCGATTTCCGCGGA 4740  
Qy 4741 GTCCCTCTTCTCATGTCAACGCTGGGTACAGGAGTCTGCGGGCGCAGCGCATCATG 4800  
Db 4741 GTCCCTCTTCTCATGTCAACGCTGGGTACAGGAGTCTGCGGGCGCAGCGCATCATG 4800  
Qy 4801 CAAACCACTGCCCCATGTGAGACACAGATCACCGACATGTGAAACACGTTTCCATGAGG 4860  
Db 4801 CAAACCACTGCCCCATGTGAGACACAGATCACCGACATGTGAAACACGTTTCCATGAGG 4860  
Qy 4861 ATCGTGGGCTTAGACCTGTATGTAACAGTGGCTGGAACATTTCCCATTAACGCGTAC 4920  
Db 4861 ATCGTGGGCTTAGACCTGTATGTAACAGTGGCTGGAACATTTCCCATTAACGCGTAC 4920  
Qy 4921 ACCAGGGCCCTGACGCGCTCCCGCGCCAAATTTCTAGGCGCTGTGCGCGGTG 4980  
Db 4921 ACCAGGGCCCTGACGCGCTCCCGCGCCAAATTTCTAGGCGCTGTGCGCGGTG 4980  
Qy 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGATG 5040  
Db 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGATG 5040  
Qy 5041 ACCACTGACAAAGTAAAGTCCCGCTCCCGCGCCAAATTTCTTACAGAGTG 5100  
Db 5041 ACCACTGACAAAGTAAAGTCCCGCTCCCGCGCCAAATTTCTTACAGAGTG 5100  
Qy 5101 GATGGGTGGGTGACAGGTACGCTCCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
Db 5101 GATGGGTGGGTGACAGGTACGCTCCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
Qy 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTTGGGTCAAGCTCCATCGAGGCCGAA 5220  
Db 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTTGGGTCAAGCTCCATCGAGGCCGAA 5220  
Qy 5221 CCGGAGCTAGAGTGTCTTCCATGCTCACCGACCCCTCCACATTAACGGGAGAGC 5280  
Db 5221 CCGGAGCTAGAGTGTCTTCCATGCTCACCGACCCCTCCACATTAACGGGAGAGC 5280  
Qy 5281 GCTAAGCTAGGCTGCGCAGGAGTCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Db 5281 GCTAAGCTAGGCTGCGCAGGAGTCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340





Db 661 TGGTATTGGCGAAGTCGCGGGCAGGATCTCCCTGTCTATCTCACTTGTCTCTCGCGAGA 720  
 Qy 721 AAGTATCGATCATGGCTGATGCAATCGCGGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
 Db 721 AAGTATCGATCATGGCTGATGCAATCGCGGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
 Qy 781 CATTCGACCAACGAGCAATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC 840  
 Db 781 CATTCGACCAACGAGCAATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC 840  
 Qy 841 TTGTCGATCAGGATGATCTGCAAGAGAGATCAGGGGCTCGGCCAGCGCAATGTTTCG 900  
 Db 841 TTGTCGATCAGGATGATCTGCAAGAGAGATCAGGGGCTCGGCCAGCGCAATGTTTCG 900  
 Qy 901 CCAGGCTCAAGGCGCGATGCCGACGAGCATCAGGGGCTCGGCCAGCGCAATGTTTCG 960  
 Db 901 CCAGGCTCAAGGCGCGATGCCGACGAGCATCAGGGGCTCGGCCAGCGCAATGTTTCG 960  
 Qy 961 GCTTGGCGAATATCATGCTGGAATGCGCGCTTTCTGGAATTCATCGACTGTCGCGG 1020  
 Db 961 GCTTGGCGAATATCATGCTGGAATGCGCGCTTTCTGGAATTCATCGACTGTCGCGG 1020  
 Qy 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080  
 Db 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080  
 Qy 1081 TTGGCGGCAATGGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTCGC 1140  
 Db 1081 TTGGCGGCAATGGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTCGC 1140  
 Qy 1141 AGCGATCGCTTCTATCGCTTCTTGACAGATCTCTGAGTTTAAACAGACCAACG 1200  
 Db 1141 AGCGATCGCTTCTATCGCTTCTTGACAGATCTCTGAGTTTAAACAGACCAACG 1200  
 Qy 1201 GTTTCCTCTAGCGGGATCAATTCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGGC 1260  
 Db 1201 GTTTCCTCTAGCGGGATCAATTCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGGC 1260  
 Qy 1261 CGAAGCGCTTGAATTAAGGCGGCTGTGCTTTGTATATGTTATTTTCCACCATATTG 1320  
 Db 1261 CGAAGCGCTTGAATTAAGGCGGCTGTGCTTTGTATATGTTATTTTCCACCATATTG 1320  
 Qy 1321 CGCTCTTTTGGCAATGTAGGCGCCCGGAACCTCGCGCTCTCTTTGACGACATTCCT 1380  
 Db 1321 CGCTCTTTTGGCAATGTAGGCGCCCGGAACCTCGCGCTCTCTTTGACGACATTCCT 1380  
 Qy 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTGAAATGTCGTGAAGGAAGCA 1440  
 Db 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTGAAATGTCGTGAAGGAAGCA 1440  
 Qy 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGCAAGCAGCG 1500  
 Db 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGCAAGCAGCG 1500  
 Qy 1501 AACCCGCCACCTCGGACAGGTGCTCTCGGCCCAAGGCAAGCCAGCTGATAAGATACCT 1560  
 Db 1501 AACCCGCCACCTCGGACAGGTGCTCTCGGCCCAAGGCAAGCCAGCTGATAAGATACCT 1560  
 Qy 1561 GCAAGGCGGCAACCCAGTGCACAGTGTGTGAGTTGATGTTGTTGGAAGAGTCAAA 1620  
 Db 1561 GCAAGGCGGCAACCCAGTGCACAGTGTGTGAGTTGATGTTGTTGGAAGAGTCAAA 1620  
 Qy 1621 TGGCTCTCTCAAGCTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
 Db 1621 TGGCTCTCTCAAGCTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
 Qy 1681 ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTCGAGGTTAAA 1740  
 Db 1681 ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTCGAGGTTAAA 1740  
 Qy 1741 AACGCTTAGGCGCCCGCAACCAAGGAGCGTGGTTCCTTTTGAAGAAACAGGATATACC 1800

Db 1741 AACGCTTAGGCGCCCGCAACCAAGGAGCGTGGTTCCTTTGAAAAACAGATATACC 1800  
 Qy 1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCTTACTTGGTGTGATCATCACT 1860  
 Db 1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCTTACTTGGTGTGATCATCACT 1860  
 Qy 1861 AGCTCACAGCGCGGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACGCA 1920  
 Db 1861 AGCTCACAGCGCGGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACGCA 1920  
 Qy 1921 ACACAATCTTTCTGGGACCTCGTCAATGCGGTGTGTTGGATGTCTATCATGTGTC 1980  
 Db 1921 ACACAATCTTTCTGGGACCTCGTCAATGCGGTGTGTTGGATGTCTATCATGTGTC 1980  
 Qy 1981 GGCTCAAGAACCTTTCGCGGCCCAAGGGGCCCAATCACCCCAATGTACCAATGTGAC 2040  
 Db 1981 GGCTCAAGAACCTTTCGCGGCCCAAGGGGCCCAATCACCCCAATGTACCAATGTGAC 2040  
 Qy 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCATGCACTGC 2100  
 Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCATGCACTGC 2100  
 Qy 2101 GGCAGCTCGGACCTTTTACTTGTGTCAGAGGATGCGATGTCATTCCGGTGGCGCGCG 2160  
 Db 2101 GGCAGCTCGGACCTTTTACTTGTGTCAGAGGATGCGATGTCATTCCGGTGGCGCGCG 2160  
 Qy 2161 GGCAGCAGCAGGGGAGGCTTACTCTCCCCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
 Db 2161 GGCAGCAGCAGGGGAGGCTTACTCTCCCCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
 Qy 2221 GCGGCTTCCACTGCTCTGCCCCCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGC 2280  
 Db 2221 GCGGCTTCCACTGCTCTGCCCCCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGC 2280  
 Qy 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTGGAGTCTATGNAACCACTATG 2340  
 Db 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTGGAGTCTATGNAACCACTATG 2340  
 Qy 2341 CGGTCCCGGCTTCTCAAGGCAACTCGTCCCTCCCGCGCTACCGCAGACATTTCCAGGTG 2400  
 Db 2341 CGGTCCCGGCTTCTCAAGGCAACTCGTCCCTCCCGCGCTACCGCAGACATTTCCAGGTG 2400  
 Qy 2401 GCCCATCTACGCCCTTCTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGGTATGCA 2460  
 Db 2401 GCCCATCTACGCCCTTCTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGGTATGCA 2460  
 Qy 2461 GCCCAAGGGTATAAGGTGCTTGTCCCTGAACCGCTCGGTGCGCGCCACCTTAGGTTTCGG 2520  
 Db 2461 GCCCAAGGGTATAAGGTGCTTGTCCCTGAACCGCTCGGTGCGCGCCACCTTAGGTTTCGG 2520  
 Qy 2521 GCGTATATGTTAAGSCACATGTTATCGACCCCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
 Db 2521 GCGTATATGTTAAGSCACATGTTATCGACCCCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
 Qy 2581 ACCAAGGGTCCCGCATACGTTACTTCCACTATGGAAGTTCCTTCGCGACGGTGTTC 2640  
 Db 2581 ACCAAGGGTCCCGCATACGTTACTTCCACTATGGAAGTTCCTTCGCGACGGTGTTC 2640  
 Qy 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCCACTCACTGACTGACCACT 2700  
 Db 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCCACTCACTGACTGACCACT 2700  
 Qy 2701 ATCTGGGCGCATCGGACAGTCTCTGGAACCAAGCGGAGAGCGGCTGGAGCGGCTCGTCGTG 2760  
 Db 2701 ATCTGGGCGCATCGGACAGTCTCTGGAACCAAGCGGAGAGCGGCTGGAGCGGCTCGTCGTG 2760  
 Qy 2761 CTCGCCACCGCTACGCTCCGGGATCGGTTCACCGTGGCAATCCAAACATCGAGGAGGTG 2820  
 Db 2761 CTCGCCACCGCTACGCTCCGGGATCGGTTCACCGTGGCAATCCAAACATCGAGGAGGTG 2820  
 Qy 2821 GCTCTGTCAGACCTGGAGGATCCCTTTTATGCAAGGCAATCCCATTCGAGACCATC 2880  
 Db 2821 GCTCTGTCAGACCTGGAGGATCCCTTTTATGCAAGGCAATCCCATTCGAGACCATC 2880











QY 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAGACAAATCGGCTGCT 480  
DB 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAGACAAATCGGCTGCT 480  
QY 481 CTGATCGCGCGCTGTTCCGGCTGTGAGCGAGGGCGCCGGTCTTTTTCGTCGAAGACCG 540  
DB 481 CTGATCGCGCGCTGTTCCGGCTGTGAGCGAGGGCGCCGGTCTTTTTCGTCGAAGACCG 540  
QY 541 ACCTGTCGGGTGCCCTGAATGAATGAGGACGAGGAGCGCGCTATCGTGGCTGGCCA 600  
DB 541 ACCTGTCGGGTGCCCTGAATGAATGAGGACGAGGAGCGCGCTATCGTGGCTGGCCA 600  
QY 601 CGAGCGCGCTTCCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGC 660  
DB 601 CGAGCGCGCTTCCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGC 660  
QY 661 TGCTATTGGGCGAAGTCCCGGGGAGGATCTCTGTGTCATCTCACTTTGCTCCTGCCGAGA 720  
DB 661 TGCTATTGGGCGAAGTCCCGGGGAGGATCTCTGTGTCATCTCACTTTGCTCCTGCCGAGA 720  
QY 721 AAGTATCCATATGCTGATGCAATGGCGGGCTGTCATAGCTTTGATCCGGCTACCTGCC 780  
DB 721 AAGTATCCATATGCTGATGCAATGGCGGGCTGTCATAGCTTTGATCCGGCTACCTGCC 780  
QY 781 CATTCGACCAACGAGCGAAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCCGGTC 840  
DB 781 CATTCGACCAACGAGCGAAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCCGGTC 840  
QY 841 TTGTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGGCCAGCGCAACTGTTGC 900  
DB 841 TTGTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGGCCAGCGCAACTGTTGC 900  
QY 901 CGAGGCTCAAGGGCGGATCGCGAAGAGCATCTGTCGTGACCCATGGCGATCCCT 960  
DB 901 CGAGGCTCAAGGGCGGATCGCGAAGAGCATCTGTCGTGACCCATGGCGATCCCT 960  
QY 961 GCTTTCGGAATATCATGGTGGAAATGGCGCTTTCTGGAATTCATCGATGCGCGCGC 1020  
DB 961 GCTTTCGGAATATCATGGTGGAAATGGCGCTTTCTGGAATTCATCGATGCGCGCGC 1020  
QY 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCGGTGATTTGCTGAAGAGC 1080  
DB 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCGGTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATGCGCGCTCCCGATTCGC 1140  
DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATGCGCGCTCCCGATTCGC 1140  
QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTTCCTGAGTTTAAACAGACCAACG 1200  
DB 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTTCCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCGCTAACGTTACTGGC 1260  
DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCGCTAACGTTACTGGC 1260  
QY 1261 CGAAGCGCTTGAATGAAGCGCGGTGCGTGTGCTATATGTTTCCACCATATTG 1320  
DB 1261 CGAAGCGCTTGAATGAAGCGCGGTGCGTGTGCTATATGTTTCCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGAGGGCCCGGAAACCTGGCCCTCTCTTGTGACGACATTCCT 1380  
DB 1321 CCGTCTTTTGGCAATGAGGGCCCGGAAACCTGGCCCTCTCTTGTGACGACATTCCT 1380  
QY 1381 AGGGGCTTTTCCCTCTCGCGCAAGGAATGCAAGGTCTGTGAATGTCGTGAAGGAGCA 1440  
DB 1381 AGGGGCTTTTCCCTCTCGCGCAAGGAATGCAAGGTCTGTGAATGTCGTGAAGGAGCA 1440  
QY 1441 GTTTCCTGGAAGCTTCTTGAAGACAAACAGCTGTGTAGCGACCTTTTGACGACGCGG 1500  
DB 1441 GTTTCCTGGAAGCTTCTTGAAGACAAACAGCTGTGTAGCGACCTTTTGACGACGCGG 1500

QY 1501 AACCCCCCACTGGCGACAGGTGCTCTGGGCCCCAAAGCCACCTGTATATAGATACACT 1560  
DB 1501 AACCCCCCACTGGCGACAGGTGCTCTGGGCCCCAAAGCCACCTGTATATAGATACACT 1560  
QY 1561 GCAAGCGGCGACAAACCCAGTGCACGTTGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAGCGGCGACAAACCCAGTGCACGTTGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAAGAGGTACCCCATTTG 1680  
DB 1621 TGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAAGAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTTACATGTGTTTGTGAGGTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTTACATGTGTTTGTGAGGTAAAA 1740  
QY 1741 AACGCTTAGGCGCCCGCGAACACAGGGGACGTTGTTTTCTTTGAAAAACAGATAATACC 1800  
DB 1741 AACGCTTAGGCGCCCGCGAACACAGGGGACGTTGTTTTCTTTGAAAAACAGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCGCTACTTTGGCTGCATCATCACT 1860  
DB 1801 ATGGCGCTATTACGGCTTACTCCAAACAGACGAGGCGCTACTTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCACAGCGGAGACAGAAACAGGTCGAGGGGAGGTCCAAAGTGGTCTCCACCGCA 1920  
DB 1861 AGCCTCACAGCGGAGACAGAAACAGGTCGAGGGGAGGTCCAAAGTGGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTTGGCGACCTCGTCAATGGCGTGTGTTGGATGCTCATCATGTGGC 1980  
DB 1921 ACACATCTTTCTTGGCGACCTCGTCAATGGCGTGTGTTGGATGCTCATCATGTGGC 1980  
QY 1981 GGCTCAAGACCCCTTGGCGGCCCAAGGGGCCAAATCACCCAAATGTACACCAATGTGGAC 2040  
DB 1981 GGCTCAAGACCCCTTGGCGGCCCAAGGGGCCAAATCACCCAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCATGACCTGC 2100  
DB 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCATGACCTGC 2100  
QY 2101 GGCAGCTCGACCTTTACTTTGTCAGAGGACATGCGATGTCATTCGGTGGCGCGCGG 2160  
DB 2101 GGCAGCTCGACCTTTACTTTGTCAGAGGACATGCGATGTCATTCGGTGGCGCGCGG 2160  
QY 2161 GGCGACAGAGGGGAGCTTCTCTCCCGAGCGCGTCTCTTACTTGAAGGGCTCTTCG 2220  
DB 2161 GGCGACAGAGGGGAGCTTCTCTCCCGAGCGCGTCTCTTACTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTTGGGCTGCGGTGTC 2280  
DB 2221 GCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTCTATGGAAAAACACTATG 2340  
DB 2281 ACCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTCTATGGAAAAACACTATG 2340  
QY 2341 CGGTCCCCGCTCTTACGAGCAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTG 2400  
DB 2341 CGGTCCCCGCTCTTACGAGCAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTG 2400  
QY 2401 GCCCATCTACGCCCTTACTGCTAGCGGACAGACACTAAGTGGCGGTGCGGTATGCA 2460  
DB 2401 GCCCATCTACGCCCTTACTGCTAGCGGACAGACACTAAGTGGCGGTGCGGTATGCA 2460  
QY 2461 GCCCAAGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCTTAGGTTTCGGG 2520  
DB 2461 GCCCAAGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCTTAGGTTTCGGG 2520  
QY 2521 GGTATATGCTTAAGGCACATGATCGACCTTAAACATCAGAACCGGGGTAAAGACCATC 2580  
DB 2521 GGTATATGCTTAAGGCACATGATCGACCTTAAACATCAGAACCGGGGTAAAGACCATC 2580  
QY 2581 ACCAGGGGTGCCCCCATCATCTCACTATGGCAAGTTTCTTGGCGAGCGGTGTTGC 2640

Db 2581 ACCAGGTCGCCCATACAGTACTCCACCTATGCAAGTTTCTTGGCGACGGTGGTGC 2640  
Qy TCTGGGGCGCCTATGACATATAATATGTGTAGTGCACACTCAACTGACTCGACCACT 2700  
Db TCTGGGGCGCCTATGACATATAATATGTGTAGTGCACACTCAACTGACTCGACCACT 2700  
Qy ATCTGGGGATCGGCACAGTCTCTGGACACAGCGGAGAGCGGTGGAGCGGCACTCGTGGT 2760  
Db ATCTGGGGATCGGCACAGTCTCTGGACACAGCGGAGAGCGGTGGAGCGGCACTCGTGGT 2760  
Qy CTCGCCACCGCTACCGCTCCGGGATCGGTACCGTGCCACATCCAAACATCCAGAGGTTG 2820  
Db CTCGCCACCGCTACCGCTCCGGGATCGGTACCGTGCCACATCCAAACATCCAGAGGTTG 2820  
Qy GCTCTGTCCAGACTGGGAAATCCCTTTTATGCGAAAGCCATCCCATCGAGACCATC 2880  
Db GCTCTGTCCAGACTGGGAAATCCCTTTTATGCGAAAGCCATCCCATCGAGACCATC 2880  
Qy AAGGGGGGAGGACCTCAATTTCTGCCATTCGAAGAAATGTATGAGTTCGCGCG 2940  
Db AAGGGGGGAGGACCTCAATTTCTGCCATTCGAAGAAATGTATGAGTTCGCGCG 2940  
Qy AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGATGATCCGTC 3000  
Db AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGATGATCCGTC 3000  
Qy ATACCAACTAGCGGAGACGTCATTTGTCGTAGCAACCGACGCTCTAATGACGGGCTTTAC 3060  
Db ATACCAACTAGCGGAGACGTCATTTGTCGTAGCAACCGACGCTCTAATGACGGGCTTTAC 3060  
Qy GCGGATTTTCGACTCGAGTATGACATGATGTCATGATGTCATGATGATGATGATGATG 3120  
Db GCGGATTTTCGACTCGAGTATGACATGATGTCATGATGATGATGATGATGATGATGATG 3120  
Qy CTGGACCCGACCTTACCACTTACAGACGACCGTGCCACAGACGCGGTGTACGCTCG 3180  
Db CTGGACCCGACCTTACCACTTACAGACGACCGTGCCACAGACGCGGTGTACGCTCG 3180  
Qy CAGCGCGAGGAGGACGCTGGTAGGGGACGATGGGCATTTACAGTTTGTGATCCAGGA 3240  
Db CAGCGCGAGGAGGACGCTGGTAGGGGACGATGGGCATTTACAGTTTGTGATCCAGGA 3240  
Qy GAAAGCCCTCGGGCATGTTGATTTCTCGGTTCTGTGGAGTGTATGACGCGGCTGT 3300  
Db GAAAGCCCTCGGGCATGTTGATTTCTCGGTTCTGTGGAGTGTATGACGCGGCTGT 3300  
Qy GCTTGTGACGAGCTACGCGCCGCGGAGACCTCAGTTAGGTTGCGGGCTTTACCTAAACACA 3360  
Db GCTTGTGACGAGCTACGCGCCGCGGAGACCTCAGTTAGGTTGCGGGCTTTACCTAAACACA 3360  
Qy CAGAGGTTCCCGTCTGCGCAGACGATCTGGAGTCTGGAGAGCGCTTTACAGGCTC 3420  
Db CAGAGGTTCCCGTCTGCGCAGACGATCTGGAGTCTGGAGAGCGCTTTACAGGCTC 3420  
Qy ACCCATAGACGCCATTTCTGTCACAGCTTAAGCAGCAGAGACCACTTCCCTTAC 3480  
Db ACCCATAGACGCCATTTCTGTCACAGCTTAAGCAGCAGAGACCACTTCCCTTAC 3480  
Qy CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
Db CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
Qy CAAATGTGAAGTGTCTACAGGCTTAAGCTTACGTCAGCGGCAACGCCCTGCTG 3600  
Db CAAATGTGAAGTGTCTACAGGCTTAAGCTTACGTCAGCGGCAACGCCCTGCTG 3600  
Qy TATAGCTGGAGCCGTTCAAAACAGAGTTACTACACACACACCCCAATCAATACATC 3660  
Db TATAGCTGGAGCCGTTCAAAACAGAGTTACTACACACACACCCCAATCAATACATC 3660  
Qy ATGGCATGATGTCGGCTGACCTGGAGGTCGTACGAGACCTGGGTGCTGTAGGCGGA 3720  
Db ATGGCATGATGTCGGCTGACCTGGAGGTCGTACGAGACCTGGGTGCTGTAGGCGGA 3720

Db 3661 ATGCATGATGTCGGCTGACCTGGAGGTCTGTCAGAGCACCTGGGTGCTGTAGCGGA 3720  
Qy GTCTTAGCAGCTCTGGCGCGCTATTGCTGCTGACAAAGCAGCAGCTGCTCATTTGGGCGAG 3780  
Db GTCTTAGCAGCTCTGGCGCGCTATTGCTGCTGACAAAGCAGCAGCTGCTCATTTGGGCGAG 3780  
Qy ATCTCTTTGTCGGAAAGCCGCGCATCTTTCCGACAGAGGAGTCTTTTACCGGGAGTTTC 3840  
Db ATCTCTTTGTCGGAAAGCCGCGCATCTTTCCGACAGAGGAGTCTTTTACCGGGAGTTTC 3840  
Qy GATGAGTGGAGAGTGGCGCTCAACCTCCCTTACATCGAACAGGGAATGAGCTCGCC 3900  
Db GATGAGTGGAGAGTGGCGCTCAACCTCCCTTACATCGAACAGGGAATGAGCTCGCC 3900  
Qy GAACTATTCAACAGAGGCAATCGGTTGCTGCAACAGCACCAGCAAGCAGGAGGT 3960  
Db GAACTATTCAACAGAGGCAATCGGTTGCTGCAACAGCACCAGCAAGCAGGAGGT 3960  
Qy GCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCTGAAAGCCTTCTGGGCGAAGCATATG 4020  
Db GCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCTGAAAGCCTTCTGGGCGAAGCATATG 4020  
Qy TGGAAATTTTCATCAGCGGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCC 4080  
Db TGGAAATTTTCATCAGCGGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCC 4080  
Qy GCGATAGCATCACTGATGCGATTACAGCTCTATCACCAGCCCGCTCACACCCCAACAT 4140  
Db GCGATAGCATCACTGATGCGATTACAGCTCTATCACCAGCCCGCTCACACCCCAACAT 4140  
Qy ACCCTCTCTGTTTAAACATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCACGCGT 4200  
Db ACCCTCTCTGTTTAAACATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCACGCGT 4200  
Qy GTCTGCTTTTTCGTAGGCGCCGGCATCGCTGAGCGCTGTGTGGCAGCATAGGCTTGG 4260  
Db GTCTGCTTTTTCGTAGGCGCCGGCATCGCTGAGCGCTGTGTGGCAGCATAGGCTTGG 4260  
Qy AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCTGTCTGGGC 4320  
Db AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCTGTCTGGGC 4320  
Qy TTTAAAGGTCAAGCGCGGAGATGCCCTCCACCGAGGACCTGTTTAACTTCTCTCTGCT 4380  
Db TTTAAAGGTCAAGCGCGGAGATGCCCTCCACCGAGGACCTGTTTAACTTCTCTCTGCT 4380  
Qy ATCTCTCTCCCTTGGCGCCCTAGTCTGCGGGTGTGTGGCAGCGATCTGCTGCGGCAC 4440  
Db ATCTCTCTCCCTTGGCGCCCTAGTCTGCGGGTGTGTGGCAGCGATCTGCTGCGGCAC 4440  
Qy GTGGGCCACAGGGGAGGGGCTGTGCAAGTGAATGAAACCGCTGTAGCGTTCGCTTCGCG 4500  
Db GTGGGCCACAGGGGAGGGGCTGTGCAAGTGAATGAAACCGCTGTAGCGTTCGCTTCGCG 4500  
Qy GGTAAACACGCTCTCCCGCACGCACTATGTGCTGAGAGGAGCGCTGACAGCGTGTCACT 4560  
Db GGTAAACACGCTCTCCCGCACGCACTATGTGCTGAGAGGAGCGCTGACAGCGTGTCACT 4560  
Qy CAGATCTCTCTAGCTTACCATCACTCAGCTGCTGAAAGGCTTACAGGTGATCAAC 4620  
Db CAGATCTCTCTAGCTTACCATCACTCAGCTGCTGAAAGGCTTACAGGTGATCAAC 4620  
Qy GAGGACTGTCTCAACGCGATGCTCCGGCTCGTGAAGAGTGTGGGATGGATATGC 4680  
Db GAGGACTGTCTCAACGCGATGCTCCGGCTCGTGAAGAGTGTGGGATGGATATGC 4680  
Qy ACGGTGTTGACTGATTTTAAAGACCTGCTCAGTCCAGTCTCTGCGGGAITTCGCGGGA 4740  
Db ACGGTGTTGACTGATTTTAAAGACCTGCTCAGTCCAGTCTCTGCGGGAITTCGCGGGA 4740  
Qy GTCCCCCTTCTCTCATGTCAAGTGGGTACAGGGAGTCTGGCGGGGCGACGATCATG 4800  
Db GTCCCCCTTCTCTCATGTCAAGTGGGTACAGGGAGTCTGGCGGGGCGACGATCATG 4800

4801 CAAACACCTGCGCATGTGGAGCACAGATCACCGACATGTGAAAACGGTTCCATGAGG 4860  
4801 CAAACACCTGCGCATGTGGAGCACAGATCACCGACATGTGAAAACGGTTCCATGAGG 4860  
4861 ATCGTGGGCGCTAGGACCTGTAGTAAACACGTGGCATGGAACATTTCCCATTAACCGGTAC 4920  
4861 ATCGTGGGCGCTAGGACCTGTAGTAAACACGTGGCATGGAACATTTCCCATTAACCGGTAC 4920  
4921 ACCACGGGCGCTGACGCGCTCCCGCGCGCAAAATTTCTAGGGCGCTGTGCGGGTG 4980  
4921 ACCACGGGCGCTGACGCGCTCCCGCGCGCAAAATTTCTAGGGCGCTGTGCGGGTG 4980  
4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGCGGATG 5040  
4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGCGGATG 5040  
5041 ACCACTGACAACTGTAAGTCCCGTGTACAGTTCCCGGCGCGCAAAATTTCTTACAGAAAGTG 5100  
5041 ACCACTGACAACTGTAAGTCCCGTGTACAGTTCCCGGCGCGCAAAATTTCTTACAGAAAGTG 5100  
5101 GATGGGCTGGGTGACACAGTAGCTCCAGCGTGCAGGTTGCAAAACCCCTTCACGGGAGGTC 5160  
5101 GATGGGCTGGGTGACACAGTAGCTCCAGCGTGCAGGTTGCAAAACCCCTTCACGGGAGGTC 5160  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCGCGAA 5220  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCGCGAA 5220  
5221 CCGGAGCTAGCATGCTCACTTCCATGCTCAACGACCCCTCCACATTAACGGCGGAGCG 5280  
5221 CCGGAGCTAGCATGCTCACTTCCATGCTCAACGACCCCTCCACATTAACGGCGGAGCG 5280  
5281 GCTAAGCTAGGCTGGCAGGAGATCCCGCTCTTGGCGAGCTCATCAGCTAGCCAG 5340  
5281 GCTAAGCTAGGCTGGCAGGAGATCCCGCTCTTGGCGAGCTCATCAGCTAGCCAG 5340  
5341 CTGCTCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCACTGCTCCCGGAGCTGAC 5400  
5341 CTGCTCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCACTGCTCCCGGAGCTGAC 5400  
5401 CTCAATCGAGCCCAACCTCTCTGCGGCGAGAGATGGCGGGGAACATCAACCCGCTGGAG 5460  
5401 CTCAATCGAGCCCAACCTCTCTGCGGCGAGAGATGGCGGGGAACATCAACCCGCTGGAG 5460  
5461 TCAGAAATAAGTAGTAATTTTGGACTCTTTCGAGCCGCTCCAGCGGAGGAGATGAG 5520  
5461 TCAGAAATAAGTAGTAATTTTGGACTCTTTCGAGCCGCTCCAGCGGAGGAGATGAG 5520  
5521 AGGGAAGTATCCGTTCCGGCGGAGATCCCTGCGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
5521 AGGGAAGTATCCGTTCCGGCGGAGATCCCTGCGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
5581 CCCATATGGGCGCGCCGGATTAACACCTCCACTGTAGAGTCCCTGGAAGGACCCCGGAC 5640  
5581 CCCATATGGGCGCGCCGGATTAACACCTCCACTGTAGAGTCCCTGGAAGGACCCCGGAC 5640  
5641 TAGCTCCCTCCAGTGATACAGGGTGTCCATTGCCCGCTCCAGCGCCCTCCGATACCA 5700  
5641 TAGCTCCCTCCAGTGATACAGGGTGTCCATTGCCCGCTCCAGCGCCCTCCGATACCA 5700  
5701 CCTCCACGGAGGAAGGAGCGGTTGTCTGTGAGAAATCTACCGTGTCTTCTCGCTTGGCG 5760  
5701 CCTCCACGGAGGAGGAGCGGTTGTCTGTGAGAAATCTACCGTGTCTTCTCGCTTGGCG 5760  
5761 GAGCTCGGCCACAAAGACCTTCGGAGCTCCGAATCGTGGCGGTCCGACAGCGCGACCGCA 5820  
5761 GAGCTCGGCCACAAAGACCTTCGGAGCTCCGAATCGTGGCGGTCCGACAGCGCGACCGCA 5820  
5821 ACGGCTCTCTGACACAGCCCTCCGACGAGCGGAGCGCGGATTCGGAATTCGAGTCGTAC 5880  
5821 ACGGCTCTCTGACACAGCCCTCCGACGAGCGGAGCGCGGATTCGGAATTCGAGTCGTAC 5880  
5821 ACGGCTCTCTGACACAGCCCTCCGACGAGCGGAGCGCGGATTCGGAATTCGAGTCGTAC 5880

5881 TCCTCCATGCCCCCTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGACGGGTCTTGG 5940  
5881 TCCTCCATGCCCCCTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGACGGGTCTTGG 5940  
5941 TCTACCCCTAGCGAGGAGGCTAGTGAGGACGTCTCTGCTGCTCGATGTCTTACACATGG 6000  
5941 TCTACCCGTAGCGAGGAGGCTAGTGAGGAGCTCTCTGCTGCTCGATGTCTTACACATGG 6000  
6001 ACAGGCGCCCTGATCAGCCCATGCGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
6001 ACAGGCGCCCTGATCAGCCCATGCGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
6061 AGCAACTCTTTGCTCCCTACCAAACTTGGTCTATGCTACTAAACATCTCGCAGCGCAAGC 6120  
6061 AGCAACTCTTTGCTCCCTACCAAACTTGGTCTATGCTACTAAACATCTCGCAGCGCAAGC 6120  
6121 CTGCGGAGAGAGGCTACCTTTGACAGACTGCAAGTCTTGAAGCTTAACTTCTATCCGTGGAG 6180  
6121 CTGCGGAGAGAGGCTACCTTTGACAGACTGCAAGTCTTGAAGCTTAACTTCTATCCGTGGAG 6180  
6181 GTGCTCAAGGAGATGAAGGCGAAGGGCTCCACAGTTAAGGCTTAACTTCTATCCGTGGAG 6240  
6181 GTGCTCAAGGAGATGAAGGCGAAGGGCTCCACAGTTAAGGCTTAACTTCTATCCGTGGAG 6240  
6241 GAAGCTGTAAAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6300  
6241 GAAGCTGTAAAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6300  
6301 GAGCTCGGAACCTATCCAGCAAGGCGGTTAACACCATCGCTCCGCTCGGTGGAAGGACTG 6360  
6301 GAGCTCGGAACCTATCCAGCAAGGCGGTTAACACCATCGCTCCGCTCGGTGGAAGGACTG 6360  
6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATAGGTTTCTGCG 6420  
6361 CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAATAGGTTTCTGCG 6420  
6421 GTCCACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6480  
6421 GTCCACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6480  
6481 GTTCGTGTGCGGAGAAATGGCCCTTTACGATGTGTGTCTCCACCCCTCCCTCAGGCGGTG 6540  
6481 GTTCGTGTGCGGAGAAATGGCCCTTTACGATGTGTGTCTCCACCCCTCCCTCAGGCGGTG 6540  
6541 ATGGGCTCTTCATACGGATTTCCAAATCTCTCTGGAGACAGGGGTGAGTTCCTGTGTGAAT 6600  
6541 ATGGGCTCTTCATACGGATTTCCAAATCTCTCTGGAGACAGGGGTGAGTTCCTGTGTGAAT 6600  
6601 GCTTGAAGGCGAAGAAATGCGCTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660  
6601 GCTTGAAGGCGAAGAAATGCGCTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660  
6661 ACGGCTCAGTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGTTGTGACTTGGCC 6720  
6661 ACGGCTCAGTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGTTGTGACTTGGCC 6720  
6721 CCCGAAGCCAGACAGGCGCATAGGTCGCTCAGAGCGGCTTTACATCGGGGCGCCCTG 6780  
6721 CCCGAAGCCAGACAGGCGCATAGGTCGCTCAGAGCGGCTTTACATCGGGGCGCCCTG 6780  
6781 ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGCGGCGGCTGTACTGACG 6840  
6781 ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGCGGCGGCTGTACTGACG 6840  
6841 ACCAGCTCGGTAAATACCTCAGATGTTACTTGAAGGCGGCTGCGGCGGCTGTGAGCTGGC 6900  
6841 ACCAGCTCGGTAAATACCTCAGATGTTACTTGAAGGCGGCTGCGGCGGCTGTGAGCTGGC 6900  
6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAAGC 6960  
6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAAGC 6960  
6961 GCGGGGACCCAGAGGACGAGCGGAGCTCAAGGCGCTTTCAGCGGCTATGACTAGATAC 7020

Db 6961 GCGGGGACCCAGAGGACGAGGCGAGCTACGGGCTTCACGGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTCGGGACCCGCCCAACCAAGTACGACTTGGAGTTGATTAACATCATGC 7080  
Db 7021 TCTGCCCCCTCGGGACCCGCCCAACCAAGTACGACTTGGAGTTGATTAACATCATGC 7080  
QY 7081 TCCTCCAAATGTGTCAGTCCGACGATCATCTGCGAAAGGGTGTATCTCTCACCCGT 7140  
Db 7081 TCCTCCAAATGTGTCAGTCCGACGATCATCTGCGAAAGGGTGTATCTCTCACCCGT 7140  
QY 7141 GACCCACACCCCTGCGGGGGCTGCTGGGAGACAGCTAGACACATCCAGTCAAT 7200  
Db 7141 GACCCACACCCCTGCGGGGGCTGCTGGGAGACAGCTAGACACATCCAGTCAAT 7200  
QY 7201 TCCTGGCTAGGCAACATCATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTTCTCCATCTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTGATG 7320  
Db 7261 ACTCATTTCTTCTCCATCTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTGATG 7320  
QY 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGCACCTACCTCAGATCATTCACGACTC 7380  
Db 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGCACCTACCTCAGATCATTCACGACTC 7380  
QY 7381 CATGGCTTAGCGCACTTTTCACTCCATGATTTACTCTCCAGTGTAGATCAATAGGCTGCT 7440  
Db 7381 CATGGCTTAGCGCACTTTTCACTCCATGATTTACTCTCCAGTGTAGATCAATAGGCTGCT 7440  
QY 7441 TCATGCTCAGAAACTTGGGGTACGCGCTTGGAGTCTGAGACATCGGCCGCAAGT 7500  
Db 7441 TCATGCTCAGAAACTTGGGGTACGCGCTTGGAGTCTGAGACATCGGCCGCAAGT 7500  
QY 7501 GTCGCGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCGCGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGAGTAGGACCAAGCTCAAACTCACTCCAATCCCGCTGCGTCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAGGACCAAGCTCAAACTCACTCCAATCCCGCTGCGTCCAGTTGGAT 7620  
QY 7621 TTATCAGCTGGTTGCTGCTGTATGCTGCTACTCTACTCTTCTGTAGGGGTAGGCATCTAT 7680  
Db 7621 TTATCAGCTGGTTGCTGCTGTATGCTGCTACTCTACTCTTCTGTAGGGGTAGGCATCTAT 7680  
QY 7681 GCGCGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7740  
Db 7681 GCGCGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7740  
QY 7741 CTACTCCCAACCGATGAAGGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCAACCGATGAAGGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
QY 7801 TTTCCCTTT 7860  
Db 7801 TTTCCCTTT 7860  
QY 7861 TTTTCTCTCTTT 7920  
Db 7861 TTTTCTCTCTTT 7920  
QY 7921 TAGCTGTGAAGCTCCGTGAGCCGCTGTACTGAGAGAGTGTGTACTGGCCCTCTCTGC 7980  
Db 7921 TAGCTGTGAAGCTCCGTGAGCCGCTGTACTGAGAGAGTGTGTACTGGCCCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

ABK91412  
ID ABK91412 standard; DNA; 10690 BP.  
XX  
AC ABK91412;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus vector construct pHCVNeo.17.  
XX  
KW HCV; ss; pHCVNeo.17; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRBS; NS5A; HCV replication.  
XX  
OS Hepatitis C virus.  
OS Encephalomyocarditis virus.  
OS Escherichia coli.  
OS Enterobacter phage T7.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
FT 5'UTR 1..341  
FT /tag= a  
FT CDS 342..1181  
FT /tag= b  
FT /product= "Core-neo fusion protein"  
FT misc\_signal 1190..1800  
FT /tag= c  
FT /label= IRES  
FT /note= "Internal ribosome entry site from EMCV"  
FT CDS 1801..7758  
FT /tag= d  
FT /product= "Polyprotein"  
FT /note= "Comprising NS3, NS4A, NS4B, NS5A and NS5B"  
FT 3'UTR 7759..7989  
FT /tag= e  
FT misc\_feature 7990..10690  
FT /tag= f  
FT /note= "Plasmid derived sequences"  
XX  
WO200259321-A2.  
XX  
XX 01-AUG-2002.  
XX  
XX 16-JAN-2002; 2002WO-EP000526.  
XX  
XX 23-JAN-2001; 2001US-0263479P.  
XX  
XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
XX  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
XX ribosome entry site (IRES) region, useful in studying HCV replication and  
XX expression.  
XX  
XX Claim 1; Fig 1; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
XX internal ribosome entry site (IRES) region coding for one or more NS3,  
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
XX are detailed in the specification. Also included are (1) an expression  
XX vector comprising a nucleotide sequence coding for the altered nucleic  
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
XX recombinant cell human hepatoma cell comprising the altered nucleic acids  
XX; (3) a recombinant cell produced by introducing an HCV (hepatitis C virus)  
XX cell the altered nucleic acids; (4) producing an HCV replicon; (5)  
XX an HCV replicon enhanced cells made in the method; and (6) measuring the  
XX ability of a compound to affect HCV activity. The HCV replicons and HCV

CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is the HCV based vector  
CC pHCNeo.17, used as a basis for the adaptive mutations of the invention  
XX  
SQ Sequence 10690 BP; 2334 A; 3045 C; 2908 G; 2403 T; 0 U; 0 Other;

Query Match 100.0%; Score 7987.4; DB 6; Length 10690;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GCCAGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTAGT	60
DB	1	GCCAGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTAGT	60
QY	61	TCCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCCTCCAGGAC	120
DB	61	TCCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCCTCCAGGAC	120
QY	121	CCCCCTCCCGGAGAGCATTAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGAG	180
DB	121	CCCCCTCCCGGAGAGCATTAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGAG	180
QY	181	GAGACCGGGTCTCTTCTTGATCAACCCGCTCAATGCTGGAGATTTGGCGTGGCCCC	240
DB	181	GAGACCGGGTCTCTTCTTGATCAACCCGCTCAATGCTGGAGATTTGGCGTGGCCCC	240
QY	241	GCGAGACTGTAGCCGAGTGTGTGGTCCGAAAGCGCTTGTGTACTGCTGTATAGG	300
DB	241	GCGAGACTGTAGCCGAGTGTGTGGTCCGAAAGCGCTTGTGTACTGCTGTATAGG	300
QY	301	GTCTTTCGAGTGGCCCGGAGGTTCTGTAGACCGTGCACCATGAGCAAGAACTCAAC	360
DB	301	GTCTTTCGAGTGGCCCGGAGGTTCTGTAGACCGTGCACCATGAGCAAGAACTCAAC	360
QY	361	CTCAAGAAACCAAGAGCGCGCATGATTGAAACAGATGATTCACGCGAGTTCTC	420
DB	361	CTCAAGAAACCAAGAGCGCGCATGATTGAAACAGATGATTCACGCGAGTTCTC	420
QY	421	CGCGCGCTTGGTGGAGAGCTATTGGCTATGACTGGGCAACAGACAAATCGGTGCT	480
DB	421	CGCGCGCTTGGTGGAGAGCTATTGGCTATGACTGGGCAACAGACAAATCGGTGCT	480
QY	481	CTGATCGCGCGTTCGGCTGTGAGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540
DB	481	CTGATCGCGCGTTCGGCTGTGAGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540
QY	541	ACCTGTCCGCTGCTGATGAATGAATGAACTGACGAGCGAGCGCGGCTATCGTGGCTG	600
DB	541	ACCTGTCCGCTGCTGATGAATGAATGAACTGACGAGCGAGCGCGGCTATCGTGGCTG	600
QY	601	CGACGGGCTTCTTTCGCGAGCTGTCTGCAAGTGTCTGCAAGCGGGAAGGAGTGGC	660
DB	601	CGACGGGCTTCTTTCGCGAGCTGTCTGCAAGTGTCTGCAAGCGGGAAGGAGTGGC	660
QY	661	TGCTATTGGGCAAGTGGCGGGGAGGATCTCTGTCTCTCACTTGTCTGCGGAGA	720
DB	661	TGCTATTGGGCAAGTGGCGGGGAGGATCTCTGTCTCTCACTTGTCTGCGGAGA	720
QY	721	AAGTATCCATCATGGCTGATGAATGCAATGCGGCGGCTGATACGTTTACCTGCTG	780
DB	721	AAGTATCCATCATGGCTGATGAATGCAATGCGGCGGCTGATACGTTTACCTGCTG	780
QY	781	CATTGACCAACGAGCAACATCGCATCGAGCGAGCAGTACTCGATGGAAGCGGTC	840
DB	781	CATTGACCAACGAGCAACATCGCATCGAGCGAGCAGTACTCGATGGAAGCGGTC	840
QY	841	TTGTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGGAGCGAATGTTGG	900
DB	841	TTGTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGGAGCGAATGTTGG	900

QY	901	CCAGGCTCAAGCGCGCATGCCCGAGCGAGGATCTCTGTGACCCATGCGGATGCT	960
DB	901	CCAGGCTCAAGCGCGCATGCCCGAGCGAGGATCTCTGTGACCCATGCGGATGCT	960
QY	961	GCTTTCGGAATATCATGTGTGAAATGCGCGCTTTTCTGGATTCATCGACTGTGCGCGC	1020
DB	961	GCTTTCGGAATATCATGTGTGAAATGCGCGCTTTTCTGGATTCATCGACTGTGCGCGC	1020
QY	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCGTGGGTACCGGTGATATGCTGAAGAC	1080
DB	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCGTGGGTACCGGTGATATGCTGAAGAC	1080
QY	1081	TTGCGCGGAATGCGCTTCCCTGCTGCTTTACGCTATCGCGTATCGCGCTCCCGATTG	1140
DB	1081	TTGCGCGGAATGCGCTTCCCTGCTGCTTTACGCTATCGCGTATCGCGCTCCCGATTG	1140
QY	1141	AGCGCATGCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG	1200
DB	1141	AGCGCATGCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG	1200
QY	1201	GTTTCCCTTAGCGGATCAATTCGCGCTTCCCTGCTGCTTCCCGCTTAAAGTCTAGTGC	1260
DB	1201	GTTTCCCTTAGCGGATCAATTCGCGCTTCCCTGCTGCTTCCCGCTTAAAGTCTAGTGC	1260
QY	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCGTCTTCTATATGTTTATTTCCACCATATG	1320
DB	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCGTCTTCTATATGTTTATTTCCACCATATG	1320
QY	1321	CCGTCTTTTGGAAATGAGGCGCGGAAACCTTGGCCCTCTTCTTTCGAGCAATTCCT	1380
DB	1321	CCGTCTTTTGGCAATGAGGCGCGGAAACCTTGGCCCTCTTCTTTCGAGCAATTCCT	1380
QY	1381	AGGGCTCTTTCCTCTCGCGCAAGGAATCGAAGTCTGTGATGTCGTGAGGAAGCA	1440
DB	1381	AGGGCTCTTTCCTCTCGCGCAAGGAATCGAAGTCTGTGATGTCGTGAGGAAGCA	1440
QY	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACCTCTGTAGCGACCCCTTTCGAGCGAGCGG	1500
DB	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACCTCTGTAGCGACCCCTTTCGAGCGAGCGG	1500
QY	1501	AACCCCGACCTGCGAGAGTCTCTGCGGCAAAAGCCAGTGTATGAAGATACACCT	1560
DB	1501	AACCCCGACCTGCGAGAGTCTCTGCGGCAAAAGCCAGTGTATGAAGATACACCT	1560
QY	1561	GCAAGCGCGCAACACCCAGTGTGAGTGTGATGTCGTGAGGAAGTCAAA	1620
DB	1561	GCAAGCGCGCAACACCCAGTGTGAGTGTGATGTCGTGAGGAAGTCAAA	1620
QY	1621	TGGCTCTCTCAAGCGTATTAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTG	1680
DB	1621	TGGCTCTCTCAAGCGTATTAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTG	1680
QY	1681	ATGGGATCTGATCTGGGCTCGGTGCAATGTTTACATGTTTGTGAGGTTTAAA	1740
DB	1681	ATGGGATCTGATCTGGGCTCGGTGCAATGTTTACATGTTTGTGAGGTTTAAA	1740
QY	1741	AACGCTCTAGGCGCGGACCGGAGCGTGGTCTTCTTGAAGAAACACCATATAC	1800
DB	1741	AACGCTCTAGGCGCGGACCGGAGCGTGGTCTTCTTGAAGAAACACCATATAC	1800
QY	1801	ATGGCGCTTATTCAGCGCTTCTCCCAACAGCGAGCGAGGCTACTTGGCTGATCATCT	1860
DB	1801	ATGGCGCTTATTCAGCGCTTCTCCCAACAGCGAGCGAGGCTACTTGGCTGATCATCT	1860
QY	1861	AGCTCTCAGCGCGGACAGGAACAGGTCTGAGGGGAGGTCAGAGTGTCTCCACGCA	1920
DB	1861	AGCTCTCAGCGCGGACAGGAACAGGTCTGAGGGGAGGTCAGAGTGTCTCCACGCA	1920
QY	1921	ACCAATCTTCTTGGCGACCTGGTCAATGCGGTGTGAGTGTCTATCATGTGCTG	1980
DB	1921	ACCAATCTTCTTGGCGACCTGGTCAATGCGGTGTGAGTGTCTATCATGTGCTG	1980











FT mat\_peptide 3697..3858  
FT /tag= f  
FT /note= "Non-structural protein 4A"  
FT mat\_peptide 3859..4641  
FT /tag= g  
FT /note= "Non-structural protein 4B"  
FT mat\_peptide 4642..5982  
FT /tag= h  
FT /note= "Non-structural protein 5A"  
FT mat\_peptide 5983..7755  
FT /tag= i  
FT /note= "Non-structural protein 5B and RNA-dependent RNA  
FT polymerase"  
FT 3' UTR 7759..7989  
FT /tag= j  
FT /note= "HCV 3' UTR"  
XX  
XX US2003045568-A1.  
XX  
XX 06-MAR-2003.  
XX  
XX 19-APR-2002; 2002US-00125940.  
XX  
XX 20-APR-2001; 2001US-0285104P.  
XX  
XX (ALTA//) ALTAMURA S.  
XX (KOCH//) KOCH U.  
XX  
XX Altamura S, Koch U;  
XX WPI; 2003-447715/42.  
XX  
XX Treating infections by hepatitis C virus and its related conditions  
XX comprises administering thiosemicarbazone compounds.  
XX  
XX Example 2; Page 9-14; 25pp; English.  
XX  
XX The invention relates to a method of treating or preventing infection by  
XX hepatitis C virus or its related conditions by delaying the onset and  
XX inhibiting replication of hepatitis C virus which comprises administering  
XX thiosemicarbazone compounds. The method is useful for treating or  
XX preventing infection by hepatitis C virus or its related conditions e.g.  
XX liver inflammation, liver failure or cirrhosis, delaying the onset and  
XX inhibiting replication of hepatitis C. The present sequence represents  
XX the hepatitis C virus expression plasmid pHCVNeo17.wt DNA  
XX  
XX Sequence 10690 BP; 2334 A; 3045 C; 2908 G; 2403 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 7987.4; DB 9; Length 10690;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 1 GCCAGCCCCGATTGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
XX 1 GCCAGCCCCGATTGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
XX  
XX 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGAGCTCCAGGAC 120  
XX 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGAGCTCCAGGAC 120  
XX  
XX 121 CCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180  
XX 121 CCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180  
XX  
XX 181 GACGACCGGTCCTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC 240  
XX 181 GACGACCGGTCCTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC 240  
XX  
XX 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGCGCTTGTGTAATGCTGCTATAGG 300  
XX 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGCGCTTGTGTAATGCTGCTATAGG 300  
XX  
XX 301 GTGCTTGCGAGTGCCCGGAGGCTCTCGTAGACCGGTGACCATGAGCAGCAATCCTAAAC 360

Db 301 GTGCTTGCGAGTGCCCGGAGGTCCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
QY 361 CTCAAAGAAAACCAAGGGCGCCCATGATTGAAACAAGATTTGACGAGGTTCTC 420  
Db 361 CTCAAAGAAAACCAAGGGCGCCCATGATTGAAACAAGATTTGACGAGGTTCTC 420  
QY 421 CCGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAATCGCGTGCT 480  
Db 421 CCGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAATCGCGTGCT 480  
QY 481 CTGATGCGCGCTGTTCCGGTTCAGCGCAAGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGCTGTTCCGGTTCAGCGCAAGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCCGGTGCCTGAATGAACTGACGAGGAGGAGCGCGGCTATCGTGCTGGCCA 600  
Db 541 ACCTGTCCGGTGCCTGAATGAACTGACGAGGAGGAGCGCGGCTATCGTGCTGGCCA 600  
QY 601 CGAGCGGCTTCTTTCGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGAGGAGCTGCG 660  
Db 601 CGAGCGGCTTCTTTCGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGAGGAGCTGCG 660  
QY 661 TGCTATTGGGCAAGTCCGGGCGAGGATCTCTGTCTATCTCACCTTCTCTCGCCGAGA 720  
Db 661 TGCTATTGGGCAAGTCCGGGCGAGGATCTCTGTCTATCTCACCTTCTCTCGCCGAGA 720  
QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCGGCTACCTGCC 780  
Db 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCGGCTACCTGCC 780  
QY 781 CATTTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTCTCGGATGCAAGCCGGTC 840  
Db 781 CATTTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTCTCGGATGCAAGCCGGTC 840  
QY 841 TTGTCGATCAGGATGATCTGCAAGAGCATCAGGGGCTCGCCGACCGCAATGTTTCG 900  
Db 841 TTGTCGATCAGGATGATCTGCAAGAGCATCAGGGGCTCGCCGACCGCAATGTTTCG 900  
QY 901 CCAGGCTCAAGGCGCGATCCGCGAGGATCTCTGTGTCGTCGACCATGCGGATGCTCCT 960  
Db 901 CCAGGCTCAAGGCGCGATCCGCGAGGATCTCTGTGTCGTCGACCATGCGGATGCTCCT 960  
QY 961 GCTTTCGCAATATCATGCTGCAAAATGCGCGCTTTCTGATTCATCATGCTGCGCGGC 1020  
Db 961 GCTTTCGCAATATCATGCTGCAAAATGCGCGCTTTCTGATTCATCATGCTGCGCGGC 1020  
QY 1021 TGGGTGCGCGGACCGCTATCAGGATAGCGTTGGCTACCCGCTGATTTGCTGAAGAGC 1080  
Db 1021 TGGGTGCGCGGACCGCTATCAGGATAGCGTTGGCTACCCGCTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGCGAAATGGGCTGACCGCTTCCCTGCTGTTTACGGTATCGCGGCTCCCGATTGCG 1140  
Db 1081 TTGGCGCGAAATGGGCTGACCGCTTCCCTGCTGTTTACGGTATCGCGGCTCCCGATTGCG 1140  
QY 1141 AGCGCATCCCTTCTATCGCTTCTTTCGAGTTCCTTCTGAGTTTAAACAGACCAACG 1200  
Db 1141 AGCGCATCCCTTCTATCGCTTCTTTCGAGTTCCTTCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACCTTACTGGC 1260  
Db 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACCTTACTGGC 1260  
QY 1261 CGAAGCGCTTGAATAAGGCGGCTGCTATGTTGCTATATGTTTTCACCATATTG 1320  
Db 1261 CGAAGCGCTTGAATAAGGCGGCTGCTATGTTGCTATATGTTTTCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGTGAGGCGCCGAAACCTGCGCCCTGCTTCTTGTGACGAGATTCCT 1380  
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCCGAAACCTGCGCCCTGCTTCTTGTGACGAGATTCCT 1380  
QY 1381 AGGGTCTTTTCCCTCTCGCAAAAGGAATCAAGGTCTGTTGAAATGCTGAGGAAGCA 1440

Db	1381	AGGGTCTTTCCCTCTGCGCAAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGCA	1440
Qy	1441	GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGCTCTGTAGCGACCTTTTGAGGAGCGG	1500
Db	1441	GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGCTCTGTAGCGACCTTTTGAGGAGCGG	1500
Qy	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAGCCACGCTGTATAGATACACT	1560
Db	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAGCCACGCTGTATAGATACACT	1560
Qy	1561	GCAAGGGCGACAAACCCAGTGCACAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGGGCGACAAACCCAGTGCACAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAAAGGGCTGAAGGATGCCGCAAGGTACCCATTGT	1680
Db	1621	TGGCTCTCTCAAGCGTATTCAACAAAGGGCTGAAGGATGCCGCAAGGTACCCATTGT	1680
Qy	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAA	1740
Db	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAA	1740
Qy	1741	ACGCTCTAGGCCCCCGAACCAACCGGACGTGGTTTTCCTTTGAAAAACACGATATACC	1800
Db	1741	ACGCTCTAGGCCCCCGAACCAACCGGACGTGGTTTTCCTTTGAAAAACACGATATACC	1800
Qy	1801	ATGGGCTCTATTACGGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCT	1860
Db	1801	ATGGGCTCTATTACGGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCT	1860
Qy	1861	AGCCTCACAGCGCGGACAGGACAGGTGCGGGGGAGTCCAAAGTGTCTCCACGCA	1920
Db	1861	AGCCTCACAGCGCGGACAGGACAGGTGCGGGGGAGTCCAAAGTGTCTCCACGCA	1920
Qy	1921	ACCAATCTTCTGCGACCTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGGC	1980
Db	1921	ACCAATCTTCTGCGACCTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGGC	1980
Qy	1981	GGCTCAAGACCTTGC CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC	2040
Db	1981	GGCTCAAGACCTTGC CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC	2040
Qy	2041	CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCCCTTGAACCATGACCTGC	2100
Db	2041	CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCCCTTGAACCATGACCTGC	2100
Qy	2101	GGCAGCTCGACCTTTACTTGGTCAACGAGCATGCCGATGTCAFTCCGGTGC CGCGCGG	2160
Db	2101	GGCAGCTCGACCTTTACTTGGTCAACGAGCATGCCGATGTCAFTCCGGTGC CGCGCGG	2160
Qy	2161	GGCGACAGCGGGAGCCTACTCTCCCGCAGCGCGCTCTCTTAATTGAAGGGCTCTTCG	2220
Db	2161	GGCGACAGCGGGAGCCTACTCTCCCGCAGCGCGCTCTCTTAATTGAAGGGCTCTTCG	2220
Qy	2221	GGGGTCCACTGCTCTCGCTCGGACAGCTGCTGGGCACTTCTGGGCTCGCGTGTGC	2280
Db	2221	GGGGTCCACTGCTCTCGCTCGGACAGCTGCTGGGCACTTCTGGGCTCGCGTGTGC	2280
Qy	2281	ACCCGAGGGTTCGGAAGGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACCACTATG	2340
Db	2281	ACCCGAGGGTTCGGAAGGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACCACTATG	2340
Qy	2341	CGGTCCCGGTCTTTCAGGCAAACTGTCCTCCCTCGGCGGTACCGCAGACATCCAGGTG	2400
Db	2341	CGGTCCCGGTCTTTCAGGCAAACTGTCCTCCCTCGGCGGTACCGCAGACATCCAGGTG	2400
Qy	2401	GCCCACTACAGCCCTTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA	2460
Db	2401	GCCCACTACAGCCCTTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA	2460
Qy	2461	GCCCAAGGGTATAAGGTGCTTCTGTAACCGGTCCGTCGCGCCACCCCTAGTTTCGGG	2520
Db	2461	GCCCAAGGGTATAAGGTGCTTCTGTAACCGGTCCGTCGCGCCACCCCTAGTTTCGGG	2520
Qy	2521	GCCTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAAACGGGGTAAAGCAATC	2580
Db	2521	GCCTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAAACGGGGTAAAGCAATC	2580
Qy	2581	ACCAAGGGTCCCGCATCAGCTACTCCACCTATGCGCAAGTTCCTTGGCGAGCGTGTGC	2640
Db	2581	ACCAAGGGTCCCGCATCAGCTACTCCACCTATGCGCAAGTTCCTTGGCGAGCGTGTGC	2640
Qy	2641	TCTGGGGCGCTATGACATCATATATGATGCTGCACTCAACTCACTGACTCGACCACT	2700
Db	2641	TCTGGGGCGCTATGACATCATATATGATGCTGCACTCAACTCACTGACTCGACCACT	2700
Qy	2701	ATCTGGGCATCGGCACAGTCTCTGGACCAAGCGGACAGCGCTGAGCGGCTCTCGTG	2760
Db	2701	ATCTGGGCATCGGCACAGTCTCTGGACCAAGCGGACAGCGCTGAGCGGCTCTCGTG	2760
Qy	2761	CTCGCCACCGCTTACGCTCCGGATCGGTACCGTGCACATCCAAACATCCAGGAGGTG	2820
Db	2761	CTCGCCACCGCTTACGCTCCGGATCGGTACCGTGCACATCCAAACATCCAGGAGGTG	2820
Qy	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACATC	2880
Db	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACATC	2880
Qy	2881	AAAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAGAAATGTGATGAGTCCGCGG	2940
Db	2881	AAAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAGAAATGTGATGAGTCCGCGG	2940
Qy	2941	AAAGTCTCGGGCTCGGACTCAATGCTGAGCAATTTACCGGGCCCTTGTATCGCTC	3000
Db	2941	AAAGTCTCGGGCTCGGACTCAATGCTGAGCAATTTACCGGGCCCTTGTATCGCTC	3000
Qy	3001	ATACCAACTAGGAGGACAGCTCATTTGCTAGCAACGAGCGCTCTAATGACGGGCTTACC	3060
Db	3001	ATACCAACTAGGAGGACAGCTCATTTGCTAGCAACGAGCGCTCTAATGACGGGCTTACC	3060
Qy	3061	GGCGATTTTCGACTCAGTGTGATCGACTCAATGCTGATCGCTCAATGACGGGCTTACC	3120
Db	3061	GGCGATTTTCGACTCAGTGTGATCGACTCAATGCTGATCGCTCAATGACGGGCTTACC	3120
Qy	3121	CTGGAACCGGCTTACCATTTGAGAGGACCGCTGCGCAAGACGGGTGTCACGCTCG	3180
Db	3121	CTGGAACCGGCTTACCATTTGAGAGGACCGCTGCGCAAGACGGGTGTCACGCTCG	3180
Qy	3181	CAGCGGCGGAGGACGAGTGTGAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGG	3240
Db	3181	CAGCGGCGGAGGACGAGTGTGAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGG	3240
Qy	3241	GAACGGGCTTCCGGCATTTGATTCCTCGGTTCTGTGCGAGTCTATGACGGGCTGT	3300
Db	3241	GAACGGGCTTCCGGCATTTGATTCCTCGGTTCTGTGCGAGTCTATGACGGGCTGT	3300
Qy	3301	GCTTGTGATGAGCTCAAGCGCGGAGACCTCAGTTAGTTTGGGGTCTTACCTAAACACA	3360
Db	3301	GCTTGTGATGAGCTCAAGCGCGGAGACCTCAGTTAGTTTGGGGTCTTACCTAAACACA	3360
Qy	3361	CCAGGTTGCGCGTCTGCCAGGACCATCTGAGTTCTGGGAGAGCGTCTTTACAGGCTC	3420
Db	3361	CCAGGTTGCGCGTCTGCCAGGACCATCTGAGTTCTGGGAGAGCGTCTTTACAGGCTC	3420
Qy	3421	ACCCACATAGAGCGGCTTCTTGTCCAGACTAAGCAGGAGGAGCACTTCCCTTAC	3480
Db	3421	ACCCACATAGAGCGGCTTCTTGTCCAGACTAAGCAGGAGGAGCACTTCCCTTAC	3480
Qy	3481	CTGTTAGCATACAGGCTTACCGTGTGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC	3540
Db	3481	CTGTTAGCATACAGGCTTACCGTGTGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC	3540
Qy	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGGGCGCAAGCCCTGCTG	3600
Db	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGGGCGCAAGCCCTGCTG	3600

QY 3601 TATAGCTGGGAGCCGTTCAAAACGAGGTTACTACACACACCCCAATACCAAAATACATC 3660  
Db 3601 TATAGCTGGGAGCCGTTCAAAACGAGGTTACTACACACACCCCAATACCAAAATACATC 3660  
QY 3661 ATGGCATCATGTGCGGTGACCTGAGGTGCTGACAGACACCTGGGTGCTGGTAGCGGA 3720  
Db 3661 ATGGCATCATGTGCGGTGACCTGAGGTGCTGACAGACACCTGGGTGCTGGTAGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCCGCGATTGTCCTGACAAACAGCAGCGGTGCTATTTGGGAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCCGCGATTGTCCTGACAAACAGCAGCGGTGCTATTTGGGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCCGCGCATCAATCCCGACAGGAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCCGCGCATCAATCCCGACAGGAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGATGCGGCTCACACCTCGCTTATCATGCAACAGGGAATCAGCTCGCC 3900  
Db 3841 GATGAGATGGAAGATGCGGCTCACACCTCGCTTATCATGCAACAGGGAATCAGCTCGCC 3900  
QY 3901 GAAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCAGCCACCAAGCAGCGAGGCT 3960  
Db 3901 GAAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCAGCCACCAAGCAGCGAGGCT 3960  
QY 3961 GCTGCTCCGCTGGTGAATCAAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGAATCAAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATACATATTTAGCAGGCTTGCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACATATTTAGCAGGCTTGCCACTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCGCGCTCACACCCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCGCGCTCACACCCACAT 4140  
QY 4141 ACCCTCTGTTTAACTCTCTGGGGGAGTGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAACTCTCTGGGGGAGTGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTCGTAGCGCGCGCATCTGAGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGCGCGCGCATCTGAGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
QY 4261 AAGGCTCTGTGATATTTTGGCAGGTTATGGAGCAGGAGGCTGGCAGCGGCTGCTGGG 4320  
Db 4261 AAGGCTCTGTGATATTTTGGCAGGTTATGGAGCAGGAGGCTGGCAGCGGCTGCTGGG 4320  
QY 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAACGAGGACCTGGTTAACTTCTCCCTGCT 4380  
Db 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAACGAGGACCTGGTTAACTTCTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGGCGCCTAGTCTGCGGGGTGCTGTGGCAGCGATATGCGGTGGCAC 4440  
Db 4381 ATCTCTCTCCCTGGGCGCCTAGTCTGCGGGGTGCTGTGGCAGCGATATGCGGTGGCAC 4440  
QY 4441 GTGGGCCCAAGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCGCGG 4500  
Db 4441 GTGGGCCCAAGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCGCGG 4500  
QY 4501 GGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCAGCGCTGCAGCAGTGTCACT 4560  
Db 4501 GGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCAGCGCTGCAGCAGTGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGGAAGAGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGGAAGAGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTGCTCCACGCCATGCTCCGGCTGCTGGCTGAAGAGATGTTGGATTGGATATGC 4680  
Db 4621 GAGGACTGCTCCACGCCATGCTCCGGCTGCTGGCTGAAGAGATGTTGGATTGGATATGC 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGATTGCGGGA 4740

Db 4681 ACGGTGTTGACTGATTTCAAGACCTGGTCCAGTCCAAAGTCTCTGCGCGATTGCGGGA 4740  
QY 4741 GTCCCTCTTCTCTCATGTCAACGTGGGTACAGGAGGTCTGCGGGGCGACCGCATCATG 4800  
Db 4741 GTCCCTCTTCTCTCATGTCAACGTGGGTACAGGAGGTCTGCGGGGCGACCGCATCATG 4800  
QY 4801 CAAACACCTCCCATGTGGAGCACAGATCACCGGACATGTGAATAAAGCTTCCATGAGG 4860  
Db 4801 CAAACACCTCCCATGTGGAGCACAGATCACCGGACATGTGAATAAAGCTTCCATGAGG 4860  
QY 4861 ATCTGGGGCCCTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCGCTAC 4920  
Db 4861 ATCTGGGGCCCTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCGCTAC 4920  
QY 4921 ACCACGGGCCCTCAGCGCCCTCCCGGCGCAATTTATTTAGGCGCTGTGGCGGTG 4980  
Db 4921 ACCACGGGCCCTCAGCGCCCTCCCGGCGCAATTTATTTAGGCGCTGTGGCGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGATTTCCCACTAGCTACCGGGCATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGATTTCCCACTAGCTACCGGGCATG 5040  
QY 5041 ACCACTGACAACTGTAAGTGCCTGTCAGTTCCGGCCCCCGCAATTTCTTACAGAAAGT 5100  
Db 5041 ACCACTGACAACTGTAAGTGCCTGTCAGTTCCGGCCCCCGCAATTTCTTACAGAAAGT 5100  
QY 5101 GATGGGTGCGGTTGACAGCTAGCTCCAGCGTGGCAACCCCTCTACGGGAGAGGTC 5160  
Db 5101 GATGGGTGCGGTTGACAGCTAGCTCCAGCGTGGCAACCCCTCTACGGGAGAGGTC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCCCGAA 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCCCGAA 5220  
QY 5221 CCGGAGTAGCAGTGTCTTCCATGCTCAACGACCCCTCCACATTTACGGCGGAGACG 5280  
Db 5221 CCGGAGTAGCAGTGTCTTCCATGCTCAACGACCCCTCCACATTTACGGCGGAGACG 5280  
QY 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTAGCCAG 5340  
QY 5341 CTGTCTGCGCCTTCTTTGAAGGCAACATGCACTACCCGTCATGACTCCCGGAGCGTGC 5400  
Db 5341 CTGTCTGCGCCTTCTTTGAAGGCAACATGCACTACCCGTCATGACTCCCGGAGCGTGC 5400  
QY 5401 CTCATCGAGGCCAACCTCTCTGTCGGGAGGATGGGCGGGAACATCAACCGCGTGGAG 5460  
Db 5401 CTCATCGAGGCCAACCTCTCTGTCGGGAGGATGGGCGGGAACATCAACCGCGTGGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGACTCTTTTCGAGCGCTCCAAAGCGAGGAGATGAG 5520  
Db 5461 TCAGAAATAAGGTAGTAAATTTTGACTCTTTTCGAGCGCTCCAAAGCGAGGAGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGCGGAGATCTGTCGGAGGTCAGGAAATTCCTCTGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGCGGAGATCTGTCGGAGGTCAGGAAATTCCTCTGAGCGATG 5580  
QY 5581 CCCATATGGGACGCGCCGATTTACAACTCTGTTAGAGTCTGGAAGGACCCCGAC 5640  
Db 5581 CCCATATGGGACGCGCCGATTTACAACTCTGTTAGAGTCTGGAAGGACCCCGAC 5640  
QY 5641 TACGTCCCTCAGTGGTACAGGGTGTCCATTTGCGGCTGCAAGGCCCTCCCGATACA 5700  
Db 5641 TACGTCCCTCAGTGGTACAGGGTGTCCATTTGCGGCTGCAAGGCCCTCCCGATACA 5700  
QY 5701 CCTCCAGGAGGAGGACGCTTGTCTGTGAGAACTCTACCGTGTCTTCTGCGCTGGG 5760  
Db 5701 CCTCCAGGAGGAGGACGCTTGTCTGTGAGAACTCTACCGTGTCTTCTGCGCTGGG 5760  
QY 5761 GAGTCTGCGCAAAAGACCTTCGGGAGTCCGAATCTGTGGCGCTCGACAGCGGACCGCA 5820





QY 7981 AGATCAAGT 7989  
 Db 7981 AGATCAAGT 7989  
 RESULT 9  
 ADC83762  
 ID ADC83762 standard; DNA; 10690 BP.  
 XX  
 AC ADC83762;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE pHVNeol7.wt plasmid containing an HCV bicistronic replicon.  
 XX  
 KW Hepatitis C virus; thiosemicarbazone;  
 KW 4-(cinnamyloxy)benzaldehyde thiosemicarbazone; RHEPLISA; Ia; hepatitis C;  
 KW HCV replication system; bicistronic RNA replicon;  
 KW neomycin phosphotransferase; human hepatoma cell line; Huh-7;  
 KW neomycin sulphate; G418; pHVNeol7.wt; replicon I377neo/NS3-3'/wt.;  
 KW hepatotropic; virucide; antiinflammatory; ds.  
 XX  
 OS Synthetic.  
 OS Unidentified.  
 OS Hepatitis C virus.  
 XX  
 FN US2003176503-A1.  
 XX  
 PD 18-SEP-2003.  
 XX  
 PF 19-APR-2002; 2002US-00125920.  
 XX  
 PR 20-APR-2001; 2001US-0285195P.  
 XX  
 PA (ALTA/) ALTAMURA S.  
 PA (KOCH/) KOCH U.  
 XX  
 PI Altamura S, Koch U;  
 XX  
 DR WPI; 2003-778475/73.  
 XX  
 PT Use of thiosemicarbazone compounds for e.g. treating and preventing  
 PT hepatitis C or its related condition, and delaying the onset of hepatitis  
 PT C or its related condition.  
 XX  
 PS Example 3; SEQ ID NO 1; 30pp; English.  
 XX  
 CC The invention discloses a method for the treatment and prevention of  
 CC hepatitis C, or its related condition, which involves the administration  
 CC of thiosemicarbazone compounds, or its salts. The inhibitory activity of  
 CC 4-(cinnamyloxy)benzaldehyde thiosemicarbazone was evaluated using  
 CC RHEPLISA assay in a 96-well microplate format. The route of  
 CC administration is oral, parenteral (e.g. subcutaneous, intravenous,  
 CC intramuscular, intrasternal injection, or infusion), by inhalation spray  
 CC or rectal. The use of the 35 compounds disclosed is specifically claimed,  
 CC e.g. 4-(cinnamyloxy)benzaldehyde thiosemicarbazone (Ia). The compounds  
 CC are useful for treating and preventing hepatitis C or its related  
 CC condition, delaying the onset of hepatitis C or its related  
 CC inhibiting replication of the hepatitis C virus. The compounds are potent  
 CC inhibitor of replication of the hepatitis C virus. The compounds are potent  
 CC systems can be obtained using various techniques. Selection of cells  
 CC capable of supporting HCV replication can be achieved using bicistronic  
 CC RNA replicons expressing a selectable marker, the neomycin  
 CC phosphotransferase. Transfection of these replicons in the human hepatoma  
 CC cell line, Huh-7, followed by cultivation in the presence of neomycin  
 CC sulphate (G418), permits the isolation of clones that support HCV  
 CC replication. The sequence presented is the pHVNeol7.wt plasmid which  
 CC contains the cDNA coding for an HCV bicistronic replicon identical to  
 CC replicon I377neo/NS3-3'/wt.  
 XX  
 SQ Sequence 10690 BP; 2334 A; 3045 C; 2908 G; 2403 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 7987.4; DB 10; Length 10690;

Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCCAGCCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Db 1 GCCAGCCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 QY 61 TCTTACGCGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTCTGAGTGTCTGAGGACTCCAGGAC 120  
 Db 61 TCTTACGCGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTCTGAGTGTCTGAGGACTCCAGGAC 120  
 QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCAG 180  
 Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCAG 180  
 QY 181 GACGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTCGAGATTGGGGGTGCCCCC 240  
 Db 181 GACGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTCGAGATTGGGGGTGCCCCC 240  
 QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGCTACTGCTCTGATAGG 300  
 Db 241 GCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGCTACTGCTCTGATAGG 300  
 QY 301 GTGCTTGGAGTCCCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGAAATCCTAAAC 360  
 Db 301 GTGCTTGGAGTCCCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGAAATCCTAAAC 360  
 QY 361 CTCAAAGAAACCAAGGCGCGCCATGATTGAACAAGATGGATTGACGAGTTCTC 420  
 Db 361 CTCAAAGAAACCAAGGCGCGCCATGATTGAACAAGATGGATTGACGAGTTCTC 420  
 QY 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGATGGATTGACGAGTTCTC 480  
 Db 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGATGGATTGACGAGTTCTC 480  
 QY 481 CTGATGCGCGCTGTTCCGGTGTGAGCGAGGGCGCCCCGGTCTCTTTTCTCAAGACCG 540  
 Db 481 CTGATGCGCGCTGTTCCGGTGTGAGCGAGGGCGCCCCGGTCTCTTTTCTCAAGACCG 540  
 QY 541 ACCTGTCCGGTCCCTGAATGACTGAGCAGGAGGAGCGCGCTATCGTGGCTGGCCA 600  
 Db 541 ACCTGTCCGGTCCCTGAATGACTGAGCAGGAGGAGCGCGCTATCGTGGCTGGCCA 600  
 QY 601 CGAGCGCGTTCCTTCCGCGAGCTGTGTCGACGTTGTCACGAGCGGAGGAGGACTGGC 660  
 Db 601 CGAGCGCGTTCCTTCCGCGAGCTGTGTCGACGTTGTCACGAGCGGAGGAGGACTGGC 660  
 QY 661 TGCTATTGGCGGAGTCCCGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCCGCGAGA 720  
 Db 661 TGCTATTGGCGGAGTCCCGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCCGCGAGA 720  
 QY 721 AAGTATCCATCATGGCTGATGCAATGGCGGCTGATACGCTTGTATCCGGGTACCTGCC 780  
 Db 721 AAGTATCCATCATGGCTGATGCAATGGCGGCTGATACGCTTGTATCCGGGTACCTGCC 780  
 QY 781 CATTCACCAACCAAGCGAAACATCGCATCGAGGAGCAGCTACTCGGATGGAAGCGGTC 840  
 Db 781 CATTCACCAACCAAGCGAAACATCGCATCGAGGAGCAGCTACTCGGATGGAAGCGGTC 840  
 QY 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTCG 900  
 Db 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTCG 900  
 QY 901 CCAGGCTCAAGCGCGCATGCGCGAGAGATCTCGTGACCCATGCGGATGCGCT 960  
 Db 901 CCAGGCTCAAGCGCGCATGCGCGAGAGATCTCGTGACCCATGCGGATGCGCT 960  
 QY 961 GCTTCCGAATATCATGGTGGAAATGSCCGCTTTTCTGGATTTCATCGACTGCGCGCG 1020  
 Db 961 GCTTCCGAATATCATGGTGGAAATGSCCGCTTTTCTGGATTTCATCGACTGCGCGCG 1020  
 QY 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGC 1080





QY 3241 GAACGCCCTCGGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
Db 3241 GAACGCCCTCGGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
QY 3301 GCTTGTACGAGCTACGCGCCGCGAGACCTCAGTTAGTTGCGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTACGAGCTACGCGCCGCGAGACCTCAGTTAGTTGCGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCGCTCTGCCAGGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTTC 3420  
Db 3361 CCAGGTTGCCGCTCTGCCAGGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTTGCCAGACTAAGCAGCAGGAGACAACTTCCCTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTTGCCAGACTAAGCAGCAGGAGACAACTTCCCTAC 3480  
QY 3481 CTGGTAGCATACGAGCTACGCTGCGGCTGCGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGGTAGCATACGAGCTACGCTGCGGCTGCGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACGCGGCCAACCGCCCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACGCGGCCAACCGCCCTGCTG 3600  
QY 3601 TATAGGCTGGAGCGGTTCAAAACGAGTTTACTACACACACCCCAACCAATACATC 3660  
Db 3601 TATAGGCTGGAGCGGTTCAAAACGAGTTTACTACACACACCCCAACCAATACATC 3660  
QY 3661 ATGCATGTCATGCTGCGCTGAGCTGAGTGTCTGACGAGCAGCTGGTGTCTGTTAGCGGA 3720  
Db 3661 ATGCATGTCATGCTGCGCTGAGCTGAGTGTCTGACGAGCAGCTGGTGTCTGTTAGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCGCGGTATGCTGACAAACAGCAGCGGTGTCATGTTGGGAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCGCGGTATGCTGACAAACAGCAGCGGTGTCATGTTGGGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCCTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGAAGTGCCTCAGCTCCTTACATCGAAACAGGGAATGCACTCGCC 3900  
Db 3841 GATGAGATGAAGTGCCTCAGCTCCTTACATCGAAACAGGGAATGCACTCGCC 3900  
QY 3901 GAAACAATTCAGAAAGGCAATCGGGTCTGCAACACGCCACCAAGACGAGGCT 3960  
Db 3901 GAAACAATTCAGAAAGGCAATCGGGTCTGCAACACGCCACCAAGACGAGGCT 3960  
QY 3961 GCTGCTCCGCTGTGGATTCAGAGTGGCGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGTGGATTCAGAGTGGCGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCACTGTGCTGCGAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCACTGTGCTGCGAACCC 4080  
QY 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCGCTCACCACCCCAAT 4140  
Db 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCGCTCACCACCCCAAT 4140  
QY 4141 ACCCTCCTGTTTAAACATCTGGGGGATGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCCTGTTTAAACATCTGGGGGATGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTCTAGGCGCGGATCGTGTAGCGGCTGTGTCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCTAGGCGCGGATCGTGTAGCGGCTGTGTCAGCATAGGCTTGGG 4260  
QY 4261 AAGTGTCTTGTGGATTTTGGCAGTTATGGAGCAGGGTGGCAGGCGCTCGTGGCC 4320  
Db 4261 AAGTGTCTTGTGGATTTTGGCAGTTATGGAGCAGGGTGGCAGGCGCTCGTGGCC 4320  
QY 4321 TTTAAGGTCAAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380

Db 4321 TTTAAGGTCAAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGCGCGCCTAGTCTGTCGGGTCGTGTGCGAGCGATATCTGCGTCCGCAC 4440  
Db 4381 ATCTCTCTCCCTGCGCGCCTAGTCTGTCGGGTCGTGTGCGAGCGATATCTGCGTCCGCAC 4440  
QY 4441 GTGGGCCACAGGCGGAGGGGCTGTGCAGTGGATGAAACCGGCTGATAGCGTTCGTTCCGG 4500  
Db 4441 GTGGGCCACAGGCGGAGGGGCTGTGCAGTGGATGAAACCGGCTGATAGCGTTCGTTCCGG 4500  
QY 4501 GGTAAACACAGTCTCCCCCAGCACTATGTGCTGAGAGCGAGCTGACAGCTGTCACT 4560  
Db 4501 GGTAAACACAGTCTCCCCCAGCACTATGTGCTGAGAGCGAGCTGACAGCTGTCACT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAAGAGGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Db 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
QY 4681 ACAGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGCTCCTGCCGAGTTGCCGGA 4740  
Db 4681 ACAGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGCTCCTGCCGAGTTGCCGGA 4740  
QY 4741 GTCCCCCTTCTCATGTCACGCTGGGTACAGGAGTCTGGCGGGCGAGCGCATCATG 4800  
Db 4741 GTCCCCCTTCTCATGTCACGCTGGGTACAGGAGTCTGGCGGGCGAGCGCATCATG 4800  
QY 4801 CAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860  
Db 4801 CAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAACAAGTGGATGAAACATTTCCCATTTAACCGGTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAACAAGTGGATGAAACATTTCCCATTTAACCGGTAC 4920  
QY 4921 ACCACGGCCCTGACGCGCTCCCGGCGCCAAATTAATTTCTAGGGCGCTGTGGCGGGT 4980  
Db 4921 ACCACGGCCCTGACGCGCTCCCGGCGCCAAATTAATTTCTAGGGCGCTGTGGCGGGT 4980  
QY 4981 GCTGCTGAGGAGTACCTGAGGTTACGCGGTCGGGATTTCCACACTGACGAGCGCATG 5040  
Db 4981 GCTGCTGAGGAGTACCTGAGGTTACGCGGTCGGGATTTCCACACTGACGAGCGCATG 5040  
QY 5041 ACCACTGACAACTGAAAGTCCGCTGACGTTCCGGCCCCCGAAATTTCTTACAGAGTG 5100  
Db 5041 ACCACTGACAACTGAAAGTCCGCTGACGTTCCGGCCCCCGAAATTTCTTACAGAGTG 5100  
QY 5101 GATGGGTCGGGTGACAGGTAACGCTCCAGCGTGCAAAACCCCTCTTACGGGAGGTC 5160  
Db 5101 GATGGGTCGGGTGACAGGTAACGCTCCAGCGTGCAAAACCCCTCTTACGGGAGGTC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGAGCCGAA 5220  
QY 5221 CCGGAGTAGAGTGTCTCACTTCCATGCTCACCGACCCCTCCCAATTTACGGCGAGAGC 5280  
Db 5221 CCGGAGTAGAGTGTCTCACTTCCATGCTCACCGACCCCTCCCAATTTACGGCGAGAGC 5280  
QY 5281 GCTAAGCGTAGGTCGGCCAGGGATCTCCCGCTCTTGGCGAGCTCATAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGTCGGCCAGGGATCTCCCGCTCTTGGCGAGCTCATAGCTAGCCAG 5340  
QY 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCCGTCATGCTCCCGGAGCGCTGAC 5400  
Db 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCCGTCATGCTCCCGGAGCGCTGAC 5400  
QY 5401 CTCTCGAGGCAACCTCTCTGTCGGGAGGATGGCGGGAACTCACCGCGGTGAG 5460

[illegible]

QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGAGACATATATACAGCCTGTCTCCT 7680  
Db |||||  
QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGAGACATATATACAGCCTGTCTCCT 7680  
Db |||||  
QY 7681 GCCGACCCCGCTGGTTCATGTCCTACTCTACTCTTCTGTAGGGTAGGCATCTAT 7740  
Db |||||  
QY 7681 GCCGACCCCGCTGGTTCATGTCCTACTCTACTCTTCTGTAGGGTAGGCATCTAT 7740  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGTGAGCGGCTTGACTGCAGAGAGTCTGATCTGGCTCTCTGC 7980  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGTGAGCGGCTTGACTGCAGAGAGTCTGATCTGGCTCTCTGC 7980  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
RESULT 10  
ADP86271  
ID ADP86271 standard; DNA; 11313 BP.  
XX AC ADP86271;  
XX DT 23-SEP-2004 (first entry)  
XX DE Hepatitis C virus Con-1 replicon I377/NS3-3' derived plasmid DNA, pZSL0.  
XX KW Hepatitis C virus; HCV; anti-HCV agent; HCV infection; therapy; plasmid;  
XX OS Hepatitis C virus.  
XX FN WO200405216-A2.  
XX PD 01-JUL-2004.  
XX PF 12-DEC-2003; 2003WO-US039722.  
XX PR 13-DEC-2002; 2002US-0433303P.  
XX PA (FOXC-) FOX CHASE CANCER CENT.  
XX PI Zhu Q, Guo J, Seeger C;  
XX WPI; 2004-488079/46.  
XX New cell-line that replicates hepatitis C virus (HCV), where the cell  
PT line is selected from a non-human cell line and a human non-hepatic cell  
PT line, useful for identifying anti-HCV agents for treating HCV infections.  
XX Example III; SEQ ID NO 8; 130pp; English.  
XX The present invention provides hepatitis C virus (HCV) replication cells  
CC and cell lines derived from human non-hepatic cells or non-human cells.  
CC The invention is useful for identifying anti-HCV agents for treating HCV  
CC infections. The present sequence is hepatitis C virus Con-1 replicon  
CC I377/NS3-3' derived plasmid DNA.  
XX  
SQ Sequence 11313 BP; 2346 A; 3334 C; 3153 G; 2480 T; 0 U; 0 Other;  
Query Match 100.0%; Score 7987.4; DB 12; Length 11313;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCAGCCCCCGATTGGGGGAGACATCTCCACCATAGATCATCTCCCTGTGAGGAACACTG 60  
Db |||||  
QY 1 GCAGCCCCCGATTGGGGGAGACATCTCCACCATAGATCATCTCCCTGTGAGGAACACTG 60  
Db |||||  
QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATAGTGTGCTGAGAGCTTGGGGTGCAGGAC 120  
Db |||||  
QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATAGTGTGCTGAGAGCTTGGGGTGCAGGAC 120  
Db |||||  
QY 121 CCCCCCTCCCGGAGAGCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATGCCAG 180  
Db |||||  
QY 121 CCCCCCTCCCGGAGAGCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATGCCAG 180  
Db |||||  
QY 181 GACGACCGGGTCTTTCTTGGATCAACCGGTCAATGCTGAGATTTGGGGTGCAGGAC 240  
Db |||||  
QY 181 GACGACCGGGTCTTTCTTGGATCAACCGGTCAATGCTGAGATTTGGGGTGCAGGAC 240  
Db |||||  
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTGCTCATAGG 300  
Db |||||  
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTGCTCATAGG 300  
Db |||||  
QY 301 GTGCTTGGAGTGCCTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGAACTCTAAAC 360  
Db |||||  
QY 301 GTGCTTGGAGTGCCTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGAACTCTAAAC 360  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGCCCATGATTGAACAGATGGATTGACGAGGTTCTC 420  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGCCCATGATTGAACAGATGGATTGACGAGGTTCTC 420  
Db |||||  
QY 421 CGGCCGTGGGTGAGAGGCTATTGCGCTATGACTGGGACACAGACATCGGCTGCT 480  
Db |||||  
QY 421 CGGCCGTGGGTGAGAGGCTATTGCGCTATGACTGGGACACAGACATCGGCTGCT 480  
Db |||||  
QY 481 CTGATCGCGGCTGTTCCGGCTGTACGAGGGCGCCGGTCTTTTGTCAAGACG 540  
Db |||||  
QY 481 CTGATCGCGGCTGTTCCGGCTGTACGAGGGCGCCGGTCTTTTGTCAAGACG 540  
Db |||||  
QY 541 ACCTGTCGGTGCCTGAATGAATGAACTGAGCAGGAGGCGGCTATCGTGGTGGCCA 600  
Db |||||  
QY 541 ACCTGTCGGTGCCTGAATGAATGAACTGAGCAGGAGGCGGCTATCGTGGTGGCCA 600  
Db |||||  
QY 601 CGACGGCGTTCCTTGGCAGCTGTGCTGACAGCTGTGCTGACAGCTGTGCTGACAGCGGAGGACTGGC 660  
Db |||||  
QY 601 CGACGGCGTTCCTTGGCAGCTGTGCTGACAGCTGTGCTGACAGCGGAGGACTGGC 660  
Db |||||  
QY 661 TGCTATTGGCGAAGTGCCTGGGCGAGATCTCTGTCTATCTACCTTGTCTTCCCGGAGA 720  
Db |||||  
QY 661 TGCTATTGGCGAAGTGCCTGGGCGAGATCTCTGTCTATCTACCTTGTCTTCCCGGAGA 720  
Db |||||  
QY 721 AAGTATCCATCATGGCTGATGCAATGGGGGCTGATACGTTGATTCGGCTACCTGCC 780  
Db |||||  
QY 721 AAGTATCCATCATGGCTGATGCAATGGGGGCTGATACGTTGATTCGGCTACCTGCC 780  
Db |||||  
QY 781 CATTCGACCACCAAGCGAAACATCGCATCGAGCAGCAGTACTCGATGGAAGCGGTC 840  
Db |||||  
QY 781 CATTCGACCACCAAGCGAAACATCGCATCGAGCAGCAGTACTCGATGGAAGCGGTC 840  
Db |||||  
QY 841 TTGTGATCAGATGATCTGGACGAAGAGCATCAGGGGCTCGCGGCGAGCAACTGTTCCG 900  
Db |||||  
QY 841 TTGTGATCAGATGATCTGGACGAAGAGCATCAGGGGCTCGCGGCGAGCAACTGTTCCG 900  
Db |||||  
QY 901 CCAGCTCAAGCGCGCATGCCGAGCGGAGGATCTGTCGTGACCATGCGATGCT 960  
Db |||||  
QY 901 CCAGCTCAAGCGCGCATGCCGAGCGGAGGATCTGTCGTGACCATGCGATGCT 960  
Db |||||  
QY 961 GCTTCCCAATATCATGTGGAAATGCCGCTTTTCTGGATTCTACATGTCGTCGCGC 1020  
Db |||||  
QY 961 GCTTCCCAATATCATGTGGAAATGCCGCTTTTCTGGATTCTACATGTCGTCGCGC 1020  
Db |||||  
QY 1021 TGGGTGTCGCGAGCGCTATCAGGACATAGCGTTGGCTTACCGGTGATATTGCTGAAGAC 1080  
Db |||||

Db	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCGTGATATTGCTGAAGAC	1080
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCCTCGTCTTACGTTACGGTATCGCGCTCCCGATTCCG	1140
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCCTCGTCTTACGTTACGGTATCGCGCTCCCGATTCCG	1140
Qy	1141	AGCGCATCGCCTTCATCGCCTTCCTGACGAGTCTCTCTGAGTTTAAACAGACCAACAAG	1200
Db	1141	AGCGCATCGCCTTCATACGCTTCCTGACGAGTCTCTCTGAGTTTAAACAGACCAACAAG	1200
Qy	1201	GTTCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCCGCCCTTAACGTTACTGGC	1260
Db	1201	GTTCCTCTAGCGGATCAATTCGCGCCTCTCCCTCCCGCCCTTAACGTTACTGGC	1260
Qy	1261	CGAAGCGCTTGGAATAAGCGCGGTGCGTTGTCTATATGTTATTTCCACATATTG	1320
Db	1261	CGAAGCGCTTGGAATAAGCGCGGTGCGTTGTCTATATGTTATTTCCACATATTG	1320
Qy	1321	CGCTCTTTTGGCAATGTGAGGCGCCGGAACCTCGGCCCTCTCTTTCGACGAGATTCTCT	1380
Db	1321	CGCTCTTTTGGCAATGTGAGGCGCCGGAACCTCGGCCCTCTCTTTCGACGAGATTCTCT	1380
Qy	1381	AGGGGTCTTTCCCTCTCGGCCAAGGAATGCAAGTCTGTTGAATGTCGTGAAGGACGA	1440
Db	1381	AGGGGTCTTTCCCTCTCGGCCAAGGAATGCAAGTCTGTTGAATGTCGTGAAGGACGA	1440
Qy	1441	GTTCTCTGGAAGCTTCTTGAAGCAACAACGTCCTGAGCGACCTTTGACGACGCGG	1500
Db	1441	GTTCTCTGGAAGCTTCTTGAAGCAACAACGTCCTGAGCGACCTTTGACGACGCGG	1500
Qy	1501	AACCCCTTCTGCGGACAGTGCCTCTGCGGCCAAGCAACGTCGTATAGATACACCT	1560
Db	1501	AACCCCTTCTGCGGACAGTGCCTCTGCGGCCAAGCAACGTCGTATAGATACACCT	1560
Qy	1561	GCAAGCGGACAAACCCGAGTGCCAGTGTGATGATGAGTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGCGGACAAACCCGAGTGCCAGTGTGATGATGAGTGTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGGTATTCAACAGGGCTGAAGGATGCCAGAGGTACCCCATTTGT	1680
Db	1621	TGGCTCTCTCAAGGTATTCAACAGGGCTGAAGGATGCCAGAGGTACCCCATTTGT	1680
Qy	1681	ATGGGATCTGATCTGGGCTCTGGTGACATGCTTTACATGTGTTAGTCGAGGTTAAAA	1740
Db	1681	ATGGGATCTGATCTGGGCTCTGGTGACATGCTTTACATGTGTTAGTCGAGGTTAAAA	1740
Qy	1741	AACGCTTAGCGCCCGGAAACACAGCGGACGTGGTTTCTTCTTGAAAAACGATATACC	1800
Db	1741	AACGCTTAGCGCCCGGAAACACAGCGGACGTGGTTTCTTCTTGAAAAACGATATACC	1800
Qy	1801	ATGGCGCTATTAGGCGCTACTCCCAAAGCGGAGGCTACTTTGGCTGCATCATCACT	1860
Db	1801	ATGGCGCTATTAGGCGCTACTCCCAAAGCGGAGGCTACTTTGGCTGCATCATCACT	1860
Qy	1861	AGCTCTACAGCGCGGACAGGAACACAGTTCGAGGGGAGGTCCAGTGTGTTCCACGCA	1920
Db	1861	AGCTCTACAGCGCGGACAGGAACACAGTTCGAGGGGAGGTCCAGTGTGTTCCACGCA	1920
Qy	1921	ACACAAATCTTCTGGGACCTGCGTCAATGCGGTGTGTGAGTCTGTATCATGTGCG	1980
Db	1921	ACACAAATCTTCTGGGACCTGCGTCAATGCGGTGTGTGAGTCTGTATCATGTGCG	1980
Qy	1981	GGCTCAAGAACCTTTCGCGGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Db	1981	GGCTCAAGAACCTTTCGCGGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Qy	2041	CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGGTTCTTGTGACACCATGCACCTGC	2100
Db	2041	CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGGTTCTTGTGACACCATGCACCTGC	2100
Qy	2101	GGGAGCTCGGACCTTTACTTGGTCAGAGGATGCCGATGCTCATTCGGTGGCCCGCGG	2160
Db	2101	GGGAGCTCGGACCTTTACTTGGTCAGAGGATGCCGATGCTCATTCGGTGGCCCGCGG	2160

2161	QY	GGGACAGACGGGAGGAGCTACTCTCCCGCAGGCCGCTCTCTACTATCGAAGGGCTCTTCG	2220
2161	DB	GGGACAGACGGGAGGAGGCTACTCTCCCGCAGGCCGCTCTCTACTATCGAAGGGCTCTTCG	2220
2221	QY	GGCGTCCACTGCTCTGCCCTCTGGGGCAGCGTGTGGGCACTTTTCGGGCTGCCGTGTGC	2280
2221	DB	GGCGTCCACTGCTCTGCCCTCTGGGGCAGCGTGTGGGCACTTTTCGGGCTGCCGTGTGC	2280
2281	QY	ACCGAGGGGTTCGAAAGCGGCTGACATTGTACCCGTCGAGTCTATGCGAAACCACTATG	2340
2281	DB	ACCGAGGGGTTCGAAAGCGGCTGAGCTTTGTACCCGTCGAGTCTATGCGAAACCACTATG	2340
2341	QY	CGGTCGCCGCTCTTACGGACAACCTCGTCCCTCTCGGCCGTACCGCAGACATTCACAGGTG	2400
2341	DB	CGGTCGCCGCTCTTACGGACAACCTCGTCCCTCTCGGCCGTACCGCAGACATTCACAGGTG	2400
2401	QY	GGCCATCTACACGCCCTACTCGTAGCGGCAAGAGCACTAAGGTGCCGCTGGTATGCA	2460
2401	DB	GGCCATCTACACGCCCTACTCGTAGCGGCAAGAGCACTAAGGTGCCGCTGGTATGCA	2460
2461	QY	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGTCGTCGCCCAACCTTAGTGTTCGGG	2520
2461	DB	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGTCGTCGCCCAACCTTAGTGTTCGGG	2520
2521	QY	GGGTATATGCTAAAGGCACATGTAATCGACCCCTAAACATCAGAACCGGGGTAAAGCACATC	2580
2521	DB	GGGTATATGCTAAAGGCACATGTAATCGACCCCTAAACATCAGAACCGGGGTAAAGCACATC	2580
2581	QY	ACACGGGTGCCCACTACGTAACCTTCTTGGCAAGTTCTTGGCCGACGGTGGTTCG	2640
2581	DB	ACACGGGTGCCCACTACGTAACCTTCTTGGCAAGTTCTTGGCCGACGGTGGTTCG	2640
2641	QY	TCGTGGGGGCGCTATGACATCAATAATGTGATGAGTGCACATCACTGACTGCACCACT	2700
2641	DB	TCGTGGGGGCGCTATGACATCAATAATGTGATGAGTGCACATCACTGACTGCACCACT	2700
2701	QY	ATCCTTGGGCATCGGCACAGTCTTGGACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG	2760
2701	DB	ATCCTTGGGCATCGGCACAGTCTTGGACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG	2760
2761	QY	CTGCCACCGTACGCTCGGACTCGGTACCGTGCACCTGCGCACATCCAAACATCGAGGAGTG	2820
2761	DB	CTGCCACCGTACGCTCGGACTCGGTACCGTGCACCTGCGCACATCCAAACATCGAGGAGTG	2820
2821	QY	GCTGTGCCAGCATGTGAGAAATCCCCCTTTATGGCAAAAGCCATCCCATCGAGACCATC	2880
2821	DB	GCTGTGCCAGCATGTGAGAAATCCCCCTTTATGGCAAAAGCCATCCCATCGAGACCATC	2880
2881	QY	AAGGGGGAGGACCTCATTTCTGCCATTCGAAGAGAAATGTGATGAGCTGCCCGCG	2940
2881	DB	AAGGGGGAGGACCTCATTTCTGCCATTCGAAGAGAAATGTGATGAGCTGCCCGCG	2940
2941	QY	AAGTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACCGGGCGCTTGATATCCGTC	3000
2941	DB	AAGTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACCGGGCGCTTGATATCCGTC	3000
3001	QY	ATACCAACTAGCGAGAGCGTCAATGTCTAGCAACGGACGCTCTAATCGCGGCTTTACC	3060
3001	DB	ATACCAACTAGCGAGAGCGTCAATGTCTAGCAACGGACGCTCTAATCGCGGCTTTACC	3060
3061	QY	GGCGATTTTCGACTAGTGTATCGACTGCAATACATGTGTACCCAGACAGTGCACCTCAGC	3120
3061	DB	GGCGATTTTCGACTAGTGTATCGACTGCAATACATGTGTACCCAGACAGTGCACCTCAGC	3120
3121	QY	CTGGACCGGACCTTACCATTGAGACGACCGGTGCCAAGACGGGTGTACGCTCG	3180
3121	DB	CTGGACCGGACCTTACCATTGAGACGACCGGTGCCAAGACGGGTGTACGCTCG	3180
3181	QY	CHGCGGCGAGGACGACTGTAGGGCGAGGATGGGCATTTACAGGTTTGTGATCCAGGA	3240
3181	DB	CAGCGCGAGGACGAGCTGTAGGGCGAGGATGGGCATTTACAGGTTTGTGATCCAGGA	3240

QY 3241 GAACGGCCCTCGGCGAGTTCGATCTCTGTCGAGTGTATGACGGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGCGAGTTCGATCTCTGTCGAGTGTATGACGGGGCTGT 3300  
QY 3301 GCTTGCTAGAGCTACGCGCCGCGAGACCTCAGTTAGTTAGTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGCTAGAGCTACGCGCCGCGAGACCTCAGTTAGTTAGTTGGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGGTTGCGGCTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCTCTTTTACAGGCCCTC 3420  
Db 3361 CCAGGGTTGCGGCTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCTCTTTTACAGGCCCTC 3420  
QY 3421 ACCACATAGACGCCCATTTCTTTGTCOCAGACTAAGCAGGAGGAGACAATTTCCCTTAC 3480  
Db 3421 ACCACATAGACGCCCATTTCTTTGTCOCAGACTAAGCAGGAGGAGACAATTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACGAGCTACGGTGTGCGCACAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Db 3481 CTGGTAGCATACGAGCTACGGTGTGCGCACAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAGGGCCCAACGCCCTCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAGGGCCCAACGCCCTCTGCTG 3600  
QY 3601 TATAGGCTGGGAGCGGTTCAAACGAGGTTACTACCAACACACACCCCAATAACCAATACATC 3660  
Db 3601 TATAGGCTGGGAGCGGTTCAAACGAGGTTACTACCAACACACACCCCAATAACCAATACATC 3660  
QY 3661 ATGGCATCATGTCGGCTGACCTGAGGTGCTCAGAGCACTTGGGTGCTGGTAGGCGGA 3720  
Db 3661 ATGGCATCATGTCGGCTGACCTGAGGTGCTCAGAGCACTTGGGTGCTGGTAGGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCTGCTCAACAGGACGCTGCTTATGTGGGAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCTGCTCAACAGGACGCTGCTTATGTGGGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGTGGCTCAGCTCAGCTTCCCTTACATGCAACAGGGAATGCACTCGCC 3900  
Db 3841 GATGAGATGGAAGTGGCTCAGCTCAGCTTCCCTTACATGCAACAGGGAATGCACTCGCC 3900  
QY 3901 GAAACAATTCAAACAGAGGCAATCGGCTGTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
Db 3901 GAAACAATTCAAACAGAGGCAATCGGCTGTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
QY 3961 GCTGCTCCGCTGGTGGAAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGGAAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGGATACAAATATTAGAGGCTTGTCCACTCTGCTGCTGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGGATACAAATATTAGAGGCTTGTCCACTCTGCTGCTGCAACCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCTCTATCAGCAGCGGCTCACCACCCCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCTCTATCAGCAGCGGCTCACCACCCCAACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTCTGAGCGCCCGCATCGCTGAGCGGCTTGTGGAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTCTGAGCGCCCGCATCGCTGAGCGGCTTGTGGAGCATAGGCTTGGG 4260  
QY 4261 AAGGTGCTTGTGATATTTTGGCAGGTATGAGCAGGAGGCTGGCAGGCGCGCTGCTGGCC 4320  
Db 4261 AAGGTGCTTGTGATATTTTGGCAGGTATGAGCAGGAGGCTGGCAGGCGCGCTGCTGGCC 4320  
QY 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACCGGAGACCTGTTAACTCTCCTGCT 4380

Db 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACGAGGACCTGTTTAACTCTCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCTGTCGCGAGATGCTGCTGCTGCT 4440  
Db 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCTGTCGCGAGATGCTGCTGCTGCT 4440  
QY 4441 GTGGGCCCAAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGCGTTTCGTTTCGCG 4500  
Db 4441 GTGGGCCCAAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGCGTTTCGTTTCGCG 4500  
QY 4501 GGTAAACCAAGTCTCCCGCCACGCACTATGTCCTGAGAGCGACGCTGCAGCATGTCAC 4560  
Db 4501 GGTAAACCAAGTCTCCCGCCACGCACTATGTCCTGAGAGCGACGCTGCAGCATGTCAC 4560  
QY 4561 CAGATCTCTCTGTCAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCTCTCTGTCAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTCTCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTTGGATATGC 4680  
Db 4621 GAGGACTCTCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTTGGATATGC 4680  
QY 4681 ACCTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCTGCGGGATTCGCGGA 4740  
Db 4681 ACCTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCTGCGGGATTCGCGGA 4740  
QY 4741 GTCCCTCTCTCTCATGTCACGCTGAGGAGTCAAGGAGTCTGCGGGGCGACGGCATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCACGCTGAGGAGTCAAGGAGTCTGCGGGGCGACGGCATCATG 4800  
QY 4801 CAAACCACTCTGAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGT 4860  
Db 4801 CAAACCACTCTGAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGT 4860  
QY 4861 ATCTGTTGGGCTTAGGACTCTGATTAACACGTTGGCATGGAACATTTCCCATTTAACGCTG 4920  
Db 4861 ATCTGTTGGGCTTAGGACTCTGATTAACACGTTGGCATGGAACATTTCCCATTTAACGCTG 4920  
QY 4921 ACCACGGCCCTCTGACGCTCTCCCGGCGCCAAATTTTCTAGGGGCTGTGCGGGGTG 4980  
Db 4921 ACCACGGCCCTCTGACGCTCTCCCGGCGCCAAATTTTCTAGGGGCTGTGCGGGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTTGAGGTTTACGCGGCTGGGGATTTCCACTACGTCAGCGGCTG 5040  
Db 4981 GCTGCTGAGGAGTACGTTGAGGTTTACGCGGCTGGGGATTTCCACTACGTCAGCGGCTG 5040  
QY 5041 ACCACTGACAACTGAAAGTCCCGCTGTCAGGTTCCGCGCCCGAAATTTCTTCAAGAGTG 5100  
Db 5041 ACCACTGACAACTGAAAGTCCCGCTGTCAGGTTCCGCGCCCGAAATTTCTTCAAGAGTG 5100  
QY 5101 GATGGGCTGGGTTGACAGGTTACGCTTCCAGCTGCAAGCTTCCCTCTTACGGGAGGTC 5160  
Db 5101 GATGGGCTGGGTTGACAGGTTACGCTTCCAGCTGCAAGCTTCCCTCTTACGGGAGGTC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGGAGCGGAA 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGGAGCGGAA 5220  
QY 5221 CCGGAGTAGAGTGTCTACTTCCATGCTCAGCGACCCCTCCACATTAACGGCGGAGAG 5280  
Db 5221 CCGGAGTAGAGTGTCTACTTCCATGCTCAGCGACCCCTCCACATTAACGGCGGAGAG 5280  
QY 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTACGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTACGCCAG 5340  
QY 5341 CTGCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCCGAGCGTGCAC 5400  
Db 5341 CTGCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCCGAGCGTGCAC 5400  
QY 5401 CTCATGAGGCCAACCTCTCTGTCGGGAGGATGGGCGGGAACATCACCGCGCTGGAG 5460



Db	5401	CTCATCGAGGCGAACCTCTCTGCGGCGAGAGATGGCGGGAACATCACCCCGTGGAG	5460	QY	6541	ATGGGCTCTTTCATACGGAATTCGAATATCTCTCTCGACAGCGGGTCTGAGTTCTCTGTTGAAT	6600
QY	5461	TCAGAAAATAGGTAGTAATTTTGGACTCTTTTCAGACGCGCTTCCAGCGGAGGAGATGAG	5520	Db	6541	ATGGGCTCTTTCATACGGAATTCGAATATCTCTCTCGACAGCGGGTCTGAGTTCTCTGTTGAAT	6600
Db	5461	TCAGAAAATAGGTAGTAATTTTGGACTCTTTTCAGACGCGCTTCCAGCGGAGGAGATGAG	5520	QY	6601	GCCTTGGAAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGCTGTTTGTGACTCA	6660
QY	5521	AGGGAATATCCGTTCCGCGGAGATCTCGGAGGTCCAGGAAATTCCTTCGAGCGATG	5580	Db	6601	GCCTTGGAAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGCTGTTTGTGACTCA	6660
Db	5521	AGGGAATATCCGTTCCGCGGAGATCTCGGAGGTCCAGGAAATTCCTTCGAGCGATG	5580	QY	6661	ACGCTCACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTTGACTTGGCC	6720
QY	5581	CCCATATGGGACGCGCGGATTTACAAACCTCCACTGTTAGATCTCTGGAAGAACCCGAC	5640	Db	6661	ACGCTCACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTTGACTTGGCC	6720
Db	5581	CCCATATGGGACGCGCGGATTTACAAACCTCCACTGTTAGATCTCTGGAAGAACCCGAC	5640	QY	6721	CCCGAAGCCAGACAGGCGCATAGGTCGCTCAAGAGCGGCTTACATCGGGGGCCCCCTG	6780
QY	5641	TACGTCCCTCCAGTGGTACAGGGTGTCCATTTGCCGCTGCCAAGGCCCTTCCGATACCA	5700	Db	6721	CCCGAAGCCAGACAGGCGCATAGGTCGCTCAAGAGCGGCTTACATCGGGGGCCCCCTG	6780
Db	5641	TACGTCCCTCCAGTGGTACAGGGTGTCCATTTGCCGCTGCCAAGGCCCTTCCGATACCA	5700	QY	6781	ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGCTGCGGCGCTGCTGCTGCTGCTG	6840
QY	5701	CTTCACGAGGAGAGAGAGCGTGTCTCTGTCAGAACTACCGTGTCTTCTGCTTGGCG	5760	Db	6781	ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGCTGCGGCGCTGCTGCTGCTGCTG	6840
Db	5701	CTTCACGAGGAGAGAGAGCGTGTCTCTGTCAGAACTACCGTGTCTTCTGCTTGGCG	5760	QY	6841	ACCAGCTGGGTAATACCTCACATGTTACTTGAAGGCGCTGCGGCTGCTGCTGCTGCTG	6900
QY	5761	GAGCTCGCACAAAGACCTTCCGCGAGCTCCGAACTCGTCCGCGCTCGACAGCGGCA	5820	Db	6841	ACCAGCTGGGTAATACCTCACATGTTACTTGAAGGCGCTGCGGCTGCTGCTGCTGCTG	6900
Db	5761	GAGCTCGCACAAAGACCTTCCGCGAGCTCCGAACTCGTCCGCGCTCGACAGCGGCA	5820	QY	6901	AAAGCTCCAGGACTGACGATGCTCGTATCGGAGAGAGCTTGTGCTTATCTGTGAAGC	6960
QY	5821	ACGGCTCTCTTCCAGAGGAGAGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG	5880	Db	6901	AAAGCTCCAGGACTGACGATGCTCGTATCGGAGAGAGCTTGTGCTTATCTGTGAAGC	6960
Db	5821	ACGGCTCTCTTCCAGAGGAGAGCGGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAG	5880	QY	6961	GGGGGACCCAGAGAGGAGCGGAGCTACCGGCTTCCAGGAGCTATGACTAGATAC	7020
QY	5881	TCTCCATGCCCCCTTCCAGAGGAGAGCGGAGTCCGAGTCCGAGTCCGAGTCCGAG	5940	Db	6961	GGGGGACCCAGAGAGGAGCGGAGCTACCGGCTTCCAGGAGCTATGACTAGATAC	7020
Db	5881	TCTCCATGCCCCCTTCCAGAGGAGAGCGGAGTCCGAGTCCGAGTCCGAGTCCGAG	5940	QY	7021	TCCTGCCCTTGGGAGCCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATCATGC	7080
QY	5941	TCTACCGTAAAGCAGAGAGTGTAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	Db	7021	TCCTGCCCTTGGGAGCCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATCATGC	7080
Db	5941	TCTACCGTAAAGCAGAGAGTGTAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	QY	7081	TCCTCCAAATGTGTACGTGCGGAGAGTGTGCGGAGGAGTGTGCTTATCTCAACCGT	7140
QY	6001	ACAGGCGCCCTGATACGCGCATCGGTGCGGAGGAAACCAAGCTGCCATCAATGACTG	6060	Db	7081	TCCTCCAAATGTGTACGTGCGGAGAGTGTGCGGAGGAGTGTGCTTATCTCAACCGT	7140
Db	6001	ACAGGCGCCCTGATACGCGCATCGGTGCGGAGGAAACCAAGCTGCCATCAATGACTG	6060	QY	7141	GACCCACCAACCCCTTGGCGGCTGCGGAGGAGTGTGCGGAGGAGTGTGCTTATCTCAACCGT	7200
QY	6061	AGCAACTCTTGTCTCGGTACCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6120	Db	7141	GACCCACCAACCCCTTGGCGGCTGCGGAGGAGTGTGCGGAGGAGTGTGCTTATCTCAACCGT	7200
Db	6061	AGCAACTCTTGTCTCGGTACCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6120	QY	7201	TCCTGGCTAGGCAACATCATCATGTAATGCGCCCAACCTTGTGGGCAAGGATGCTCTGATG	7260
QY	6121	CTGGCGAGAAAGGTCACTTTGACAGACTGCGGAGTCTGCGGAGGAGTCTGCGGAG	6180	Db	7201	TCCTGGCTAGGCAACATCATCATGTAATGCGCCCAACCTTGTGGGCAAGGATGCTCTGATG	7260
Db	6121	CTGGCGAGAAAGGTCACTTTGACAGACTGCGGAGTCTGCGGAGGAGTCTGCGGAG	6180	QY	7261	ACTCATTTCTTCCATCTCTAGCTCAGGAACTTGAAGGAGGAGTGTGCTTATCTCAACCGT	7320
QY	6181	GTGCTCAAGGAGATGAAGGCGAGGCGTCCACAGTTAAGGCTAACTTCTATCCGTTGAG	6240	Db	7261	ACTCATTTCTTCCATCTCTAGCTCAGGAACTTGAAGGAGGAGTGTGCTTATCTCAACCGT	7320
Db	6181	GTGCTCAAGGAGATGAAGGCGAGGCGTCCACAGTTAAGGCTAACTTCTATCCGTTGAG	6240	QY	7321	ATCTACGGGCGCTTACTCCATTTAGCTTACCTTACCTTACCTTACCTTACCTTACCTTAC	7380
QY	6241	GAAGCTGTAGCTGAGCGGCGGAGTCCAGGAGTCCAGGAGTCCAGGAGTCCAGGAG	6300	Db	7321	ATCTACGGGCGCTTACTCCATTTAGCTTACCTTACCTTACCTTACCTTACCTTACCTTAC	7380
Db	6241	GAAGCTGTAGCTGAGCGGCGGAGTCCAGGAGTCCAGGAGTCCAGGAGTCCAGGAG	6300	QY	7381	CATGGCTTAGCGCATTTTCTCTCATAGTTTACTTCTCCAGTGTAGATCAATAGGTTGGT	7440
QY	6301	GACGTCCGGAACCTTATCAGCAAGCGGCTTAAACCAATCCGCTCCGTTGGAAGACTG	6360	Db	7381	CATGGCTTAGCGCATTTTCTCTCATAGTTTACTTCTCCAGTGTAGATCAATAGGTTGGT	7440
Db	6301	GACGTCCGGAACCTTATCAGCAAGCGGCTTAAACCAATCCGCTCCGTTGGAAGACTG	6360	QY	7441	TCATGCCCTCAGGAACTTGGGTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGG	7500
QY	6361	CTGGAAGACACTGAGACACCAATTTGACACCAATTTGACACCAATTTGACACCAATTTG	6420	Db	7441	TCATGCCCTCAGGAACTTGGGTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGG	7500
Db	6361	CTGGAAGACACTGAGACACCAATTTGACACCAATTTGACACCAATTTGACACCAATTTG	6420	QY	7501	GTCCGCGCTTAGGCTTACTTCCAGGGGGGGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
QY	6421	GTCCAAACAGAGAGAGGCGGCGGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	6480	Db	7501	GTCCGCGCTTAGGCTTACTTCCAGGGGGGGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
Db	6421	GTCCAAACAGAGAGAGGCGGCGGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	6480	QY	7561	AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCCAGTCCAGTCCAGTCCAGTCCAGT	7620
QY	6481	GTTCGTGTGTGCGAGAAATGCGGCTTTACGATGTTGTTTCAACCTCCAGTCCAGTCCAG	6540	Db	7561	AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCCAGTCCAGTCCAGTCCAGTCCAGT	7620
Db	6481	GTTCGTGTGTGCGAGAAATGCGGCTTTACGATGTTGTTTCAACCTCCAGTCCAGTCCAG	6540				



QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGGAGACATATATACAGCGCTGCTCTGT 7680  
Db |||||  
QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGGAGACATATATACAGCGCTGCTCTGT 7680  
Db |||||  
QY 7681 GCGGACCGCGTGGTTCATGTTGGTCCCTACTCTCTTCTGTAGGGTAGGCATCTAT 7740  
Db |||||  
QY 7681 GCGGACCGCGTGGTTCATGTTGGTCCCTACTCTCTTCTGTAGGGTAGGCATCTAT 7740  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7861 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7861 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGAGCGCTTGACTGCAGAGAGTCTGATCTGGCCTCTCTGC 7980  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGAGCGCTTGACTGCAGAGAGTCTGATCTGGCCTCTCTGC 7980  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||

## RESULT 11

AA47281  
ID AA47281 standard; DNA; 7992 BP.  
XX  
AC AA47281;  
XX  
DT 30-AUG-2002 (first entry)  
XX  
DE Hepatitis C virus sub-genomic replicon recombinant clone HCV24.  
XX  
KW Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
XX virucide; hepatotropic; gene therapy; anti-viral; gene; ds.  
XX  
OS Hepatitis C virus.  
XX  
FN W0200238793-A2.  
XX  
PD 16-MAY-2002.  
XX  
PF 02-NOV-2001; 2001WO-US046350.  
XX  
PR 07-NOV-2000; 2000US-0245866P.  
XX  
PA (ANAD-) ANADYS PHARM INC.  
XX  
PI Bichko V;  
XX  
XX WPI; 2002-490082/52.  
XX  
PT Novel nucleic acid encoding replication competent recombinant hepatitis C  
PT virus genome useful for screening anti-hepatitis C virus therapeutics and  
PT for vaccine development.  
XX  
PS Claim 11; Page 70-75; 85pp; English.  
XX  
CC The present invention provides protein and coding sequences from  
CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
CC able to replicate efficiently when transfected into a susceptible cell  
CC line without reducing the growth rate of the cell line by more than 10  
CC fold. The sequences are useful for screening for anti-HCV therapeutics,  
CC for detecting antibodies to HCV in a biological sample such as blood,  
CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,  
CC for deriving authentic HCV components such as replication-complement non-  
CC infectious, replication-defective infection-component, and replication-

CC defective non-infectious HCV, in gene therapy or gene vaccination  
CC targeted to hepatic tissue for treating an animal infected or susceptible  
CC to HCV infection and for studying HCV infection and propagation. The  
CC present sequence is a clone of a fragment of the HCV genome designated  
CC HCV24

XX  
SQ Sequence 7992 BP; 1648 A; 2369 C; 2242 G; 1733 T; 0 U; 0 Other;  
Query Match 100.0%; Score 7985.8; DB 6; Length 7992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCAGCCCGGATTTGGGGCGACACTCCACCATAGATCATCCCTCTGTGAGAACTACTG 60  
Db |||||  
QY 1 GCCAGCCCGGATTTGGGGCGACACTCCACCATAGATCATCCCTCTGTGAGAACTACTG 60  
Db |||||  
QY 61 TCTTCACGCAGAAACGCTCTAGCCATGGCGTTAGTATGAGTGTCTGAGCCTCAGGAC 120  
Db |||||  
QY 61 TCTTCACGCAGAAACGCTCTAGCCATGGCGTTAGTATGAGTGTCTGAGCCTCAGGAC 120  
Db |||||  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATGGCAG 180  
Db |||||  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATGGCAG 180  
Db |||||  
QY 181 GACGACCGGTCCTTTCTTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCCCC 240  
Db |||||  
QY 181 GACGACCGGTCCTTTCTTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCCCC 240  
Db |||||  
QY 241 GCGAGACTGCTAGCGGAGTAGTGTGGTCCGAAAGACCTTGGTACTGCTGATAGG 300  
Db |||||  
QY 241 GCGAGACTGCTAGCGGAGTAGTGTGGTCCGAAAGACCTTGGTACTGCTGATAGG 300  
Db |||||  
QY 301 GTGCTTGGAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGAGCAGATCTTAAC 360  
Db |||||  
QY 301 GTGCTTGGAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGAGCAGATCTTAAC 360  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGGCCATGATTGAACAGATGGATTGCACGAGTCTC 420  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGGCCATGATTGAACAGATGGATTGCACGAGTCTC 420  
Db |||||  
QY 421 CGCGCGCTTGGTGGAGAGCTATTGCGCTATGACTGGGCAACACAAATCGGCTGCT 480  
Db |||||  
QY 421 CGCGCGCTTGGTGGAGAGCTATTGCGCTATGACTGGGCAACACAAATCGGCTGCT 480  
Db |||||  
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGCTTCTTTTGTCAAGACCG 540  
Db |||||  
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGCTTCTTTTGTCAAGACCG 540  
Db |||||  
QY 541 ACTGTCCGCTGCCCTGAATGAATGAACTGAGGACGAGCGCGCTATCGTGGTGGCCA 600  
Db |||||  
QY 541 ACTGTCCGCTGCCCTGAATGAATGAACTGAGGACGAGCGCGCTATCGTGGTGGCCA 600  
Db |||||  
QY 601 CGACGGCGCTTCTTGGCAGCTGTCTGAGCTGTCTGAGCGGGAAGGAGCTGGC 660  
Db |||||  
QY 601 CGACGGCGCTTCTTGGCAGCTGTCTGAGCTGTCTGAGCGGGAAGGAGCTGGC 660  
Db |||||  
QY 661 TGCTATTGGGCGAAGTGCCTGGGCGAGGATCTCTGTCATCTCACCTTCTCTGCGCAGA 720  
Db |||||  
QY 661 TGCTATTGGGCGAAGTGCCTGGGCGAGGATCTCTGTCATCTCACCTTCTCTGCGCAGA 720  
Db |||||  
QY 721 AAGTATCCATATGCTGATGCAATGCGGCGCTGCATACGCTTGTATCCGGTACTCTGCC 780  
Db |||||  
QY 781 CATTGCAACCAAGCGAAACATCATGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
Db |||||  
QY 781 CATTGCAACCAAGCGAAACATCATGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
Db |||||  
QY 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGCTCGCGCAGCGGAACTGTCG 900  
Db |||||  
QY 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGCTCGCGCAGCGGAACTGTCG 900  
Db |||||  
QY 901 CCAGGCTCAAGGCGCGCATGTCGCGAGGAGTCTCGTCTGACCCATGCGGATGCT 960

Db	901	CCAGGCTCAAGCGCGATCGATCGGACGAGGATCTCGTGAACCATCGCGAGTGCCT	960	Db	1981	GGCTCAAGAGCCCTTGGCGGCCAAAGGGCCAAATCACCCAAATGTACACCAATGTGGAC	2040
Qy	961	GCTTGGCGAATATCATGTGGAATAATGCGCGCTTTTCGAGATTCATCGACTGTGGCGGC	1020	Qy	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGGCGGTTCCTTTACACCAATGACACTGC	2100
Db	961	GCTTGGCGAATATCATGTGGAATAATGCGCGCTTTTCGAGATTCATCGACTGTGGCGGC	1020	Db	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGGCGGTTCCTTTACACCAATGACACTGC	2100
Qy	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGAGGATGCGGTGGCTACCGGTGATATTCGTGAAGAGC	1080	Qy	2101	GGCAGCTCGGACCTTTTACTTGGTCAAGGAGCATGCGGATGTCATTCCTGGTGGCGCGGCGG	2160
Db	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGAGGATGCGGTGGCTACCGGTGATATTCGTGAAGAGC	1080	Db	2101	GGCAGCTCGGACCTTTTACTTGGTCAAGGAGCATGCGGATGTCATTCCTGGTGGCGCGGCGG	2160
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGGTTCAGGATCGCGCTCCCGATTCGC	1140	Qy	2161	GGGACAGCAGGGGAGGCTACTCTCCCGCAGGCGCGTCTCTACTTGAAGGGCTTCCTCG	2220
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGGTTCAGGATCGCGCTCCCGATTCGC	1140	Db	2161	GGGACAGCAGGGGAGGCTACTCTCTCCCGCAGGCGCGTCTCTACTTGAAGGGCTTCCTCG	2220
Qy	1141	AGCGATCGCCCTTCTATCGCCCTTCTTGACGAGTCTCTCTGAGTTTAAACAGACCAACAG	1200	Qy	2221	GGCGTTCACATGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTTCGGGCTGCCGTGTGC	2280
Db	1141	AGCGATCGCCCTTCTATCGCCCTTCTTGACGAGTCTCTCTGAGTTTAAACAGACCAACAG	1200	Db	2221	GGCGTTCACATGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTTCGGGCTGCCGTGTGC	2280
Qy	1201	GTTCCTCTAGCGGATCAATTCGCGCCCTCTGCTCCCGCCCGCTTAACTTACTGCG	1260	Qy	2281	ACCCGAGGGGTTGCGAAGCGGGTGACCTTTTGTACCCGTGAGTCTATGGAAACCACTATG	2340
Db	1201	GTTCCTCTAGCGGATCAATTCGCGCCCTCTGCTCCCGCCCGCTTAACTTACTGCG	1260	Db	2281	ACCCGAGGGGTTGCGAAGCGGGTGACCTTTTGTACCCGTGAGTCTATGGAAACCACTATG	2340
Qy	1261	CGAAGCGCTTGAATAAGCGCGGTGCGTTGTCTATATGTTATTTTCCACCATATG	1320	Qy	2341	CGGTCCCGGTCTTCAACGACAACTCGTCCCTCGCGCGTACCGCAGACATTTCCAGGTG	2400
Db	1261	CGAAGCGCTTGAATAAGCGCGGTGCGTTGTCTATATGTTATTTTCCACCATATG	1320	Db	2341	CGGTCCCGGTCTTCAACGACAACTCGTCCCTCGCGCGTACCGCAGACATTTCCAGGTG	2400
Qy	1321	CGGTCTTTTGGAAATGAGGCGCCCGAAACCTGGCCCTGTCTTTGACGAGCATTCCT	1380	Qy	2401	GCCCATCTACACGCCCTTACTGTAGCGGCAAGAGACATTAAGTGGCGGCTGGGTATGCA	2460
Db	1321	CGGTCTTTTGGAAATGAGGCGCCCGAAACCTGGCCCTGTCTTTGACGAGCATTCCT	1380	Db	2401	GCCCATCTACACGCCCTTACTGTAGCGGCAAGAGACATTAAGTGGCGGCTGGGTATGCA	2460
Qy	1381	AGGGGTCTTTCCTCTCGCCTTGAAGGATGCAAGGCTGTGTGAATCTGTGAAGGAGCA	1440	Qy	2461	GCCAAAGGATTAAGGTGTCTGTGAACCCGTGGTGGCGGCGCACCTAGTTCGGG	2520
Db	1381	AGGGGTCTTTCCTCTCGCCTTGAAGGATGCAAGGCTGTGTGAATCTGTGAAGGAGCA	1440	Db	2461	GCCAAAGGATTAAGGTGTCTGTGAACCCGTGGTGGCGGCGCACCTAGTTCGGG	2520
Qy	1441	GTTCCTCTGGAAGCTTCTTGACACAAACCTCTGTAGCGACCTTTTCAGCGAGCGG	1500	Qy	2521	GGGTATATGTCTAAGGACATGCTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Db	1441	GTTCCTCTGGAAGCTTCTTGACACAAACCTCTGTAGCGACCTTTTCAGCGAGCGG	1500	Db	2521	GGGTATATGTCTAAGGACATGCTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Qy	1501	AACCCCTCTGAGGACAGGTGCTCTGGCGCAAAAGCCAGCTGTATAGATACACCT	1560	Qy	2581	ACCAAGGGTGGCCCTTACAGTACTCCACTTGGCAAGTTTCTTGGCGAGCGTGGTTC	2640
Db	1501	AACCCCTCTGAGGACAGGTGCTCTGGCGCAAAAGCCAGCTGTATAGATACACCT	1560	Db	2581	ACCAAGGGTGGCCCTTACAGTACTCCACTTGGCAAGTTTCTTGGCGAGCGTGGTTC	2640
Qy	1561	GCAAGCGGCAACACCCAGTCCAGCTGTGAGTTGATGATGTTGGGAAAGAGTCAA	1620	Qy	2641	TCCTGGGGGCGCTTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACCAT	2700
Db	1561	GCAAGCGGCAACACCCAGTCCAGCTGTGAGTTGATGATGTTGGGAAAGAGTCAA	1620	Db	2641	TCCTGGGGGCGCTTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACCAT	2700
Qy	1621	TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGGATGCCAAGAGGTACCCCATGT	1680	Qy	2701	ATCCTGGGCATCGGCACAGTCTTGGACCAAGCGAGACGGCTGGAGCGGACTCGTGTG	2760
Db	1621	TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGGATGCCAAGAGGTACCCCATGT	1680	Db	2701	ATCCTGGGCATCGGCACAGTCTTGGACCAAGCGAGACGGCTGGAGCGGACTCGTGTG	2760
Qy	1681	ATGGATCTCATCTGGGGCTCGGTGCACATGCTTTACATGTGTTTATGCGAGTTAAA	1740	Qy	2761	CTGCGCACCGCTACGCTCCGGGATCGGTCCCGTGCACATCCCGTGCACATCCAAACATCGAGAGGTG	2820
Db	1681	ATGGATCTCATCTGGGGCTCGGTGCACATGCTTTACATGTGTTTATGCGAGTTAAA	1740	Db	2761	CTGCGCACCGCTACGCTCCGGGATCGGTCCCGTGCACATCCCGTGCACATCCAAACATCGAGAGGTG	2820
Qy	1741	AACGTCTAGGCCCCCGAAACACAGGGGACGTGGTTTCTTTGAAACACGATAATACC	1800	Qy	2821	GCTCTGTCCAGCATCGAGAAATCCCTTTTATGGCAAGAGCAATCCCGATCGAGACCATC	2880
Db	1741	AACGTCTAGGCCCCCGAAACACAGGGGACGTGGTTTCTTTGAAACACGATAATACC	1800	Db	2821	GCTCTGTCCAGCATCGAGAAATCCCTTTTATGGCAAGAGCAATCCCGATCGAGACCATC	2880
Qy	1801	ATGGCGCTTATAGGCTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACT	1860	Qy	2881	AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAGAAATGTGATGAGCTCGCGCG	2940
Db	1801	ATGGCGCTTATAGGCTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACT	1860	Db	2881	AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAGAAATGTGATGAGCTCGCGCG	2940
Qy	1861	AGCTTCACAGGCGGGAAGGACAGGTCGAGGGGAGGTCCAAAGTGGTCTCCACCGCA	1920	Qy	2941	AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGATTCGCTC	3000
Db	1861	AGCTTCACAGGCGGGAAGGACAGGTCGAGGGGAGGTCCAAAGTGGTCTCCACCGCA	1920	Db	2941	AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGATTCGCTC	3000
Qy	1921	ACACAATCTTCTGGCGACTGGTCAATGGCGTGTGGACTGTCTATCATGCTGCC	1980	Qy	3001	ATACCAATACGCGAGAGCTCATTTGTGTAGCAACGGAGCTCTTAATGAGGGCTTTTACC	3060
Db	1921	ACACAATCTTCTGGCGACTGGTCAATGGCGTGTGGACTGTCTATCATGCTGCC	1980	Db	3001	ATACCAATACGCGAGAGCTCATTTGTGTAGCAACGGAGCTCTTAATGAGGGCTTTTACC	3060
Qy	1981	GGCTCAAGAGCCCTTGGCGGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040	Qy	3061	GGCGATTTTCGACTCAGTGCATGCAATCATGTGTACCCACGACAGTGCAGTTCAGC	3120
Db	1981	GGCTCAAGAGCCCTTGGCGGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040	Db	3061	GGCGATTTTCGACTCAGTGCATGCAATCATGTGTACCCACGACAGTGCAGTTCAGC	3120

QY 3121 CTGGACCCGACCTTCAACATTGAGACGACGACCGTGCACAAAGACGCGGTGTACAGCTCG 3180  
DB |||||  
3121 CTGGACCCGACCTTCAACATTGAGACGACGACCGTGCACAAAGACGCGGTGTACAGCTCG 3180  
QY 3181 CAGCGCGGAGGACGACCTGGTAGGGGACGAGTGGGCATTACAGGTTTGTGACTCCAGGA 3240  
DB |||||  
3181 CAGCGCGGAGGACGAGTGGTAGGGGACGAGTGGGCATTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGACATGTTCCGATTCCTCGGTTCTGTGGAGTGTCTATACGCGGGCTGT 3300  
DB |||||  
3241 GAACGGCCCTCGGGACATGTTCCGATTCCTCGGTTCTGTGGAGTGTCTATACGCGGGCTGT 3300  
QY 3301 GTTGGTACGAGCTCAACGCCCGCCGAGACCTCAGTTAGGTTTCGGGCTTTACCTTAACACA 3360  
DB |||||  
3301 GTTGGTACGAGCTCAACGCCCGCCGAGACCTCAGTTAGGTTTCGGGCTTTACCTTAACACA 3360  
QY 3361 CCAGGTTGCCGCTCTGCAGGACCATCTGGAGTCTGGGAGAGCTCTTTACAGGCTC 3420  
DB |||||  
3361 CCAGGTTGCCGCTCTGCAGGACCATCTGGAGTCTGGGAGAGCTCTTTACAGGCTC 3420  
QY 3421 ACCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGACGAGTGGGAGAGCTCTTTACAGGCTC 3480  
DB |||||  
3421 ACCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGACGAGTGGGAGAGCTCTTTACAGGCTC 3480  
QY 3481 CTGTTAGCATACGAGCTACGCTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
DB |||||  
3481 CTGTTAGCATACGAGCTACGCTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
QY 3541 CAATGTGGAGTCTCTCATACGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 3600  
DB |||||  
3541 CAATGTGGAGTCTCTCATACGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 3600  
QY 3601 TATAGGCTGGAGCGCTTCAACAGAGTCTACTACCAACAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 3660  
DB |||||  
3601 TATAGGCTGGAGCGCTTCAACAGAGTCTACTACCAACAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 3660  
QY 3661 ATGGCATGCAATGCGGTGACCTGGAGTGTCTACAGAGCACTGGGTCTGTGGAGCGGA 3720  
DB |||||  
3661 ATGGCATGCAATGCGGTGACCTGGAGTGTCTACAGAGCACTGGGTCTGTGGAGCGGA 3720  
QY 3721 GTCTAGCAGCTGTGGCGGCTATTGCTGACAAAGCAGCGGTGTCTATTGGGCGAG 3780  
DB |||||  
3721 GTCTAGCAGCTGTGGCGGCTATTGCTGACAAAGCAGCGGTGTCTATTGGGCGAG 3780  
QY 3781 ATCATCTGTCCGGAAGCGGCACTCATCTCCGACAGGAAAGTCTTTACCGGAGTTC 3840  
DB |||||  
3781 ATCATCTGTCCGGAAGCGGCACTCATCTCCGACAGGAAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGATGGAAGAGTGGCTCACACCTCCCTTACATCGAACAGGAAATGCAAGCTCGCC 3900  
DB |||||  
3841 GATGATGGAAGAGTGGCTCACACCTCCCTTACATCGAACAGGAAATGCAAGCTCGCC 3900  
QY 3901 GAACAAATCAACAGAGCAATCGGCTTGTGCAAAAGCAGCAAGCAAGCAAGCAAGCAAGCAAG 3960  
DB |||||  
3901 GAACAAATCAACAGAGCAATCGGCTTGTGCAAAAGCAGCAAGCAAGCAAGCAAGCAAGCAAG 3960  
QY 3961 GCTGCTCCGCTGGTGAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
DB |||||  
3961 GCTGCTCCGCTGGTGAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
DB |||||  
4021 TGGAAATTCATCAGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
QY 4081 GCGATAGCATCTAGTGGCATCAATATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4140  
DB |||||  
4081 GCGATAGCATCTAGTGGCATCAATATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4140  
QY 4141 ACCCTCTCTTTAAATCTCTGGGGGATGGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200  
DB |||||  
4141 ACCCTCTCTTTAAATCTCTGGGGGATGGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200

QY 4201 GCTTCTGCTTTCTGTAGGCGCGGCATCTGCTGAGAGCGCTGTGTGGCAGCATAGCCCTTGG 4260  
DB |||||  
4201 GCTTCTGCTTTCTGTAGGCGCGGCATCTGCTGAGAGCGCTGTGTGGCAGCATAGCCCTTGG 4260  
QY 4261 AAGGTGCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCTGTGGCC 4320  
DB |||||  
4261 AAGGTGCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCTGTGGCC 4320  
QY 4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT 4380  
DB |||||  
4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGCGCCTAGTCTGCGGGTCTGTGCGAGCGCATCTGCTGGGCAC 4440  
DB |||||  
4381 ATCTCTCTCCCTGGCGCCTAGTCTGCGGGTCTGTGCGAGCGATCTGCTGGTCGCGAC 4440  
QY 4441 GTGGGCCAGGGAGGGGCTGTGCGAGTGAATGAACCGGCTGATAGCGTTGCTTCGCG 4500  
DB |||||  
4441 GTGGGCCAGGGAGGGGCTGTGCGAGTGAATGAACCGGCTGATAGCGTTGCTTCGCG 4500  
QY 4501 GGTAAACACAGCTCTCCCCCAGCAGCTATGCTGCTGAGAGCGCTGACGACGCTGCTACT 4560  
DB |||||  
4501 GGTAAACACAGCTCTCCCCCAGCAGCTATGCTGCTGAGAGCGCTGACGACGCTGCTACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAAGGCTTCACAGTGGATCAAC 4620  
DB |||||  
4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAAGGCTTCACAGTGGATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCATGCTCGGCTCTGGGTCTGGGTGAAGATGTTGGATGATGTC 4680  
DB |||||  
4621 GAGGACTGCTCCAGCCATGCTCGGCTCTGGGTCTGGGTGAAGATGTTGGATGATGTC 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGTCCCTGCGCGGGA 4740  
DB |||||  
4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCCTGCGCGGGA 4740  
QY 4741 GTCCCTTCTCTCATGTCAAAGTGGGTCAAGGGAGTCTGGCGGGCGGACGATCATG 4800  
DB |||||  
4741 GTCCCTTCTCTCATGTCAAAGTGGGTCAAGGGAGTCTGGCGGGCGGACGATCATG 4800  
QY 4801 CAACACCTGCGCCATGTGGAGCAGCATCAACCGGACATGTGAAACCGTTCCATGAGG 4860  
DB |||||  
4801 CAACACCTGCGCCATGTGGAGCAGCATCAACCGGACATGTGAAACCGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTACAGTGGCATGGACATTTCCCATTAACGGGTAC 4920  
DB |||||  
4861 ATCGTGGGCTTAGGACCTGTAGTAAACGTGGCATGGACATTTCCCATTAACGGGTAC 4920  
QY 4921 ACCACGGGCGCTGACGCGCTTCCCGGCGCAAAATTTCTAGGCGCTGTGCGCGGTG 4980  
DB |||||  
4921 ACCACGGGCGCTGACGCGCTTCCCGGCGCAAAATTTCTAGGCGCTGTGCGCGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGAGGTTAGCGGGTGGGGATTTCCACTACGTGACGGGATG 5040  
DB |||||  
4981 GCTGCTGAGGAGTACGTGAGGTTAGCGGGTGGGGATTTCCACTACGTGACGGGATG 5040  
QY 5041 ACCACTGACAAAGTAAAGTGGCGGCTGAGGTTCCGGCCCCCGAATTTCTTACAGAGTG 5100  
DB |||||  
5041 ACCACTGACAAAGTAAAGTGGCGGCTGAGGTTCCGGCCCCCGAATTTCTTACAGAGTG 5100  
QY 5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
DB |||||  
5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGGAGCGGAA 5220  
DB |||||  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGGAGCGGAA 5220  
QY 5221 CCGGAGGTAGGAGTGTCTCATTCCATGCTACCGACCCCTCCCATTTACGCGGAGACG 5280  
DB |||||  
5221 CCGGAGGTAGGAGTGTCTCATTCCATGCTACCGACCCCTCCCATTTACGCGGAGACG 5280  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGGATCTCCCGCTCTTGGCGGCTCATCAGCTAGCCAG 5340





QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAGGCGCTTGTGTAAGTGCCTGTAGG 300  
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAGGCGCTTGTGTAAGTGCCTGTAGG 300  
QY 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAAGCAGTGCACCATGAGCAGCAATCCTAAAC 360  
DB 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAAGCAGTGCACCATGAGCAGCAATCCTAAAC 360  
QY 361 CTCAAGAAAAACAAAGGCGCGCATGATTAAGCAAGATGGATGTCAGCAGGTTCTC 420  
DB 361 CTCAAGAAAAACAAAGGCGCGCATGATTAAGCAAGATGGATGTCAGCAGGTTCTC 420  
QY 421 CGGCGCTTGGGTGAGAGGCTATTCGCTATGACTGGGCAACAGACAATCGGCTGCT 480  
DB 421 CGGCGCTTGGGTGAGAGGCTATTCGCTATGACTGGGCAACAGACAATCGGCTGCT 480  
QY 481 CTGATGCGCGCTGTTCGCGTGTGTCAGCGCAGGCGCGCGGTTCTTTTGTCAAGACCG 540  
DB 481 CTGATGCGCGCTGTTCGCGTGTGTCAGCGCAGGCGCGCGGTTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCCGCTGCTGATGAATGCAAGTGCAGGAGCAGCGCGCTATCGTGGCTGGCA 600  
DB 541 ACCTGTCCGCTGCTGATGAATGCAAGTGCAGGAGCAGCGCGCTATCGTGGCTGGCA 600  
QY 601 CGAGCGGCTTCCCTTGGCAGCTGTGCTGAGTGTGCTCACTGAAGCGGAGGAGTGGC 660  
DB 601 CGAGCGGCTTCCCTTGGCAGCTGTGCTGAGTGTGCTCACTGAAGCGGAGGAGTGGC 660  
QY 661 TGCTATTGGGCGAAGTGGCGGCGGAGGATCTCTGTGCTCACTGCTTGTCTTCCCGAGA 720  
DB 661 TGCTATTGGGCGAAGTGGCGGCGGAGGATCTCTGTGCTCACTGCTTGTCTTCCCGAGA 720  
QY 721 AAGTATCCATCATGCTGATGCAATGCGCGGCTGATAGCTGTGATCGGCTACCTGCC 780  
DB 721 AAGTATCCATCATGCTGATGCAATGCGCGGCTGATAGCTGTGATCGGCTACCTGCC 780  
QY 781 CATTCAGCACACNAGCGAATCATGCTGAGCGGAGCAGTACTCGGATGGAAGCGGTC 840  
DB 781 CATTCAGCACACNAGCGAATCATGCTGAGCGGAGCAGTACTCGGATGGAAGCGGTC 840  
QY 841 TTGCTGATCAGGATGATCTGGACGAGAGCATCAGGCGGCTCGCGCCAGCGCACTGTTCG 900  
DB 841 TTGCTGATCAGGATGATCTGGACGAGAGCATCAGGCGGCTCGCGCCAGCGCACTGTTCG 900  
QY 901 CCAAGGCTCAAGCGCGCATGCGCGAGAGATCTGCTGCTGAGCCATGGCGATGCTTC 960  
DB 901 CCAAGGCTCAAGCGCGCATGCGCGAGAGATCTGCTGCTGAGCCATGGCGATGCTTC 960  
QY 961 GCTTGGCGAATATCATGCTGGAATGCGCGCTTCTGGAATTCATCGACTGTGGCGGC 1020  
DB 961 GCTTGGCGAATATCATGCTGGAATGCGCGCTTCTGGAATTCATCGACTGTGGCGGC 1020  
QY 1021 TGGGTGCGGACCGCTATCAGACATAGCTGTGGCTACCGCTGATTTGCTGAAGAGC 1080  
DB 1021 TGGGTGCGGACCGCTATCAGACATAGCTGTGGCTACCGCTGATTTGCTGAAGAGC 1080  
QY 1081 TTGCGCGCAATGCGCTGACGCTTCTGCTGCTTTACGCTATCGCTGCTGCTGCTGCT 1140  
DB 1081 TTGCGCGCAATGCGCTGACGCTTCTGCTGCTTTACGCTATCGCTGCTGCTGCTGCT 1140  
QY 1141 AGCGCATGCTTCTATCGCTTCTGACGAGTTCCTGAGTTTAAACAGACCAACAG 1200  
DB 1141 AGCGCATGCTTCTATCGCTTCTGACGAGTTCCTGAGTTTAAACAGACCAACAG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAAAGTTCGCTGGC 1260  
DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAAAGTTCGCTGGC 1260  
QY 1261 CGAAGCGCTTGGAAATAGGCGGCTGCTGCTTGTATATGTTTTCACCATATG 1320  
DB 1261 CGAAGCGCTTGGAAATAGGCGGCTGCTGCTTGTATATGTTTTCACCATATG 1320  
QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTTCAGCAGCATTCCT 1380

DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTTCAGCAGCATTCCT 1380  
QY 1381 AGGGGTCTTTCCCTCTCGCCAAAGGATCAAGGCTCTGTTGAATGCTGTAAGGAAGCA 1440  
DB 1381 AGGGGTCTTTCCCTCTCGCCAAAGGATCAAGGCTCTGTTGAATGCTGTAAGGAAGCA 1440  
QY 1441 GTTCCTCTGGAAGCTTCTTGAAGAACAACAGCTGTGAGCGACCTTTGAGGAGCGG 1500  
DB 1441 GTTCCTCTGGAAGCTTCTTGAAGAACAACAGCTGTGAGCGACCTTTGAGGAGCGG 1500  
QY 1501 AACCCCCACCTGGGACAGTGTCTGCGGCAAAAGCCACGCTGTATATAGATACACT 1560  
DB 1501 AACCCCCACCTGGGACAGTGTCTGCGGCAAAAGCCACGCTGTATATAGATACACT 1560  
QY 1561 GCAAGGCGGACAAACCCAGTGTGACCTTGTAGTTGATAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAGGCGGACAAACCCAGTGTGACCTTGTAGTTGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGCGTGAAGGATGCCAGAGGTACCCATTGT 1680  
DB 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGCGTGAAGGATGCCAGAGGTACCCATTGT 1680  
QY 1681 ATGGGATCTGATCTGGGCTCGGTGACATGCTTTACATGTTTGTAGTCGAGGTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGCTCGGTGACATGCTTTACATGTTTGTAGTCGAGGTAAAA 1740  
QY 1741 AACGCTTAGGCCCCCGAACCACCGGGAGTGTTCCTTTGAAAAACACGATAATAC 1800  
DB 1741 AACGCTTAGGCCCCCGAACCACCGGGAGTGTTCCTTTGAAAAACACGATAATAC 1800  
QY 1801 ATGGGCTCTATTACGCGCTTATCCAAACAGACGAGGCTACTTGGCTGCAATCATCT 1860  
DB 1801 ATGGGCTCTATTACGCGCTTATCCAAACAGACGAGGCTACTTGGCTGCAATCATCT 1860  
QY 1861 AGCTCAGAGCGGACAGGACAGGTCGAGGCGGAGTCCAAAGTGTCTCCACCGCA 1920  
DB 1861 AGCTCAGAGCGGACAGGACAGGTCGAGGCGGAGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGCGGACCTGCGTCAATGCGGTGTGGAATCTCTATCATGTGTC 1980  
DB 1921 ACACATCTTTCTGCGGACCTGCGTCAATGCGGTGTGGAATCTCTATCATGTGTC 1980  
QY 1981 GGCTCAAGACCTTTCGCGCCCAAGGCGCAATCACCCAAATGTACCAATGTGGAC 2040  
DB 1981 GGCTCAAGACCTTTCGCGCCCAAGGCGCAATCACCCAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACCTGTGCTGGCAGCGCCCGCGGCGGTTCTTGAACCATGACCTGC 2100  
DB 2041 CAGGACCTGTGCTGGCAGCGCCCGCGGCGGTTCTTGAACCATGACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTTACGTCAGGAGCATGCGGATGTCTTCCGCTGCGCGCGG 2160  
DB 2101 GGCAGCTCGGACCTTTTACGTCAGGAGCATGCGGATGTCTTCCGCTGCGCGCGG 2160  
QY 2161 GGCAGCAGCAGGCGGAGCTTCTCTCCCGCGGCGGTTCTTGAAGGCTCTTCG 2220  
DB 2161 GGCAGCAGCAGGCGGAGCTTCTCTCCCGCGGCGGTTCTTGAAGGCTCTTCG 2220  
QY 2221 GGCAGCTCTGCTGCGGCTCGGCGAGCTGTGGGATCTTTCGCGCTGCGCTGTC 2280  
DB 2221 GGCAGCTCTGCTGCGGCTCGGCGAGCTGTGGGATCTTTCGCGCTGCGCTGTC 2280  
QY 2281 ACCGAGGCGTGGAAAGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCCACTATG 2340  
DB 2281 ACCGAGGCGTGGAAAGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCCACTATG 2340  
QY 2341 CGGTCCCGGTCTTCAAGCAACTGCTGCCCTCCGCGGCTACCGCAGCATTCAGGTG 2400  
DB 2341 CGGTCCCGGTCTTCAAGCAACTGCTGCCCTCCGCGGCTACCGCAGCATTCAGGTG 2400  
QY 2401 GCGCATCTACAGCGCTTCTGAGGCGGAGGAGCACTAAGGTGCGCGCTGCTATGCA 2460



Db 2401 GCCCATCTACACGCCCTTACTGGTAGCGGCAAGACGACTAAGGTGCGGCTGCTATGCA 2460  
QY 2461 GCCAAGGATTAAGTGTCTTCTCTGAACCCGTCGTCGCCGACCACTAGTTTCGGG 2520  
Db 2461 GCCAAGGATTAAGTGTCTTCTCTGAACCCGTCGTCGCCGACCACTAGTTTCGGG 2520  
QY 2521 GGGTATATGTCTAAGGCAATGGTATGACCCCTAAACATCAGAACCGGGGTAAGGACCATC 2580  
Db 2521 GGGTATATGTCTAAGGCAATGGTATGACCCCTAAACATCAGAACCGGGGTAAGGACCATC 2580  
QY 2581 ACCACGGGTGCCCCCATCAGCTACTCCACCTATGCAAGTTCTTGGCGAAGGTTGTC 2640  
Db 2581 ACCACGGGTGCCCCCATCAGCTACTCCACCTATGCAAGTTCTTGGCGAAGGTTGTC 2640  
QY 2641 TCTGGGGCGCTATGACATCATATATGTATGATGTCACCTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCTATGACATCATATATGTATGATGTCACCTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCATCGGACAGTCTCTGGACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG 2760  
Db 2701 ATCTGGGCATCGGACAGTCTCTGGACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG 2760  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACCGTCCACATCCAAAATCGAGAGGTG 2820  
Db 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACCGTCCACATCCAAAATCGAGAGGTG 2820  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGAGGACCTCAATTTCTGGCAATTCOAAGAAATGTGATGACTCGCGCG 2940  
Db 2881 AAGGGGGAGGACCTCAATTTCTGGCAATTCOAAGAAATGTGATGACTCGCGCG 2940  
QY 2941 AAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACGGGCGCTGTGATTCGCTC 3000  
Db 2941 AAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACGGGCGCTGTGATTCGCTC 3000  
QY 3001 ATACCACTAGCGGAGAGCTGATGCTAGCAATGCTGTACCCAGACAGTCTGACG 3060  
Db 3001 ATACCACTAGCGGAGAGCTGATGCTAGCAATGCTGTACCCAGACAGTCTGACG 3060  
QY 3061 GGGATTTGCACTAAGTATCGATGCAATGCTGTACCCAGACAGTCTGACG 3120  
Db 3061 GGGATTTGCACTAAGTATCGATGCAATGCTGTACCCAGACAGTCTGACG 3120  
QY 3121 CTGGACCGGACCTTCACTTACGATGAGACGACCGTGCACAGCGGCTGTCAACGCTG 3180  
Db 3121 CTGGACCGGACCTTCACTTACGATGAGACGACCGTGCACAGCGGCTGTCAACGCTG 3180  
QY 3181 CAGCGGAGGAGGAGTCTGTTAGGGGAGGATGCGGATTTACAGGTTTGTGACTCCAGG 3240  
Db 3181 CAGCGGAGGAGGAGTCTGTTAGGGGAGGATGCGGATTTACAGGTTTGTGACTCCAGG 3240  
QY 3241 GAACGGCCCTCGGGATGTTGATTCCTCGGTTCTGTGGAGTGTATGACGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGATGTTGATTCCTCGGTTCTGTGGAGTGTATGACGGGCTGT 3300  
QY 3301 GCTTGTAGAGCTACGCGCGCGGAGCTTCACTGAGTGTGGGCTGTGAGGCTGTGCT 3360  
Db 3301 GCTTGTAGAGCTACGCGCGCGGAGCTTCACTGAGTGTGGGCTGTGAGGCTGTGCT 3360  
QY 3361 CCAGGGTTGCCGCTGCGGAGGACCTGAGGTTCTGGAGGAGGCTGTTTACAGGCTC 3420  
Db 3361 CCAGGGTTGCCGCTGCGGAGGACCTGAGGTTCTGGAGGAGGCTGTTTACAGGCTC 3420  
QY 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTTAAGCAGGAGGAGCAACTTCCCTTAC 3480  
Db 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTTAAGCAGGAGGAGCAACTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACAGGCTACGGTGTGGCGGACGGGCTCAGGCTCCACCTCGTGGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGGCGGACGGGCTCAGGCTCCACCTCGTGGGAC 3540

QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAOGGGCAACGCCCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAOGGGCAACGCCCTGCTG 3600  
QY 3601 TATAGGCTGGGAGCGGTTCAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
Db 3601 TATAGGCTGGGAGCGGTTCAAACAGAGTTACTACACACACCCCAATACCAATACATC 3660  
QY 3661 ATGGCATGCATGCTCGGCTGACCTGGAGTCTGTCAGAGCACCTGGTGTGTTGAGCGGA 3720  
Db 3661 ATGGCATGCATGCTCGGCTGACCTGGAGTCTGTCAGAGCACCTGGTGTGTTGAGCGGA 3720  
QY 3721 GTCTTAGCAGTCTGGCGCGCTATGCTCTGACACAGGAGGTTGTTTACCGGAGTTC 3780  
Db 3721 GTCTTAGCAGTCTGGCGCGCTATGCTCTGACACAGGAGGTTGTTTACCGGAGTTC 3780  
QY 3781 ATCATCTTGTCCGAAAGCCCGCCATCATTTCCGACAGGAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCCCGCCATCATTTCCGACAGGAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTACATCCCGACAGGAGTCTTTTACCGGAGTTC 3900  
Db 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTACATCCCGACAGGAGTCTTTTACCGGAGTTC 3900  
QY 3901 GAACAATTCAAACAGAGGCAATCGGTTGTCGAAACAGGACCAACAGCAAGCGGAGGCT 3960  
Db 3901 GAACAATTCAAACAGAGGCAATCGGTTGTCGAAACAGGACCAACAGCAAGCGGAGGCT 3960  
QY 3961 GCTGCTCCGCTGGTGGAAATCGAAGTGGGACCTTCGAAAGCTTCTGGGGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGGAAATCGAAGTGGGACCTTCGAAAGCTTCTGGGGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATACATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATACAGCGCGCTCACACCCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATACAGCGCGCTCACACCCACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
QY 4201 GCTTCTGCTTCTGTCGCGGATCGCTGGAGCGGCTGTTGGAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTCTGTCGCGGATCGCTGGAGCGGCTGTTGGAGCATAGGCTTGGG 4260  
QY 4261 AAGGTCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCTCGTGGCC 4320  
Db 4261 AAGGTCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCTCGTGGCC 4320  
QY 4321 TTTAAGGTATGAGCGCGGATGCCCTCCACCGAGGACCTTGTAACTTCTCTCTGCT 4380  
Db 4321 TTTAAGGTATGAGCGCGGATGCCCTCCACCGAGGACCTTGTAACTTCTCTCTGCT 4380  
QY 4381 ATCTCTCTCCCTTGGCGGCTTGTGTCGCGGCTGTCGCGAGCATATGCGTGGGAC 4440  
Db 4381 ATCTCTCTCCCTTGGCGGCTTGTGTCGCGGCTGTCGCGAGCATATGCGTGGGAC 4440  
QY 4441 GTGGGCCAGGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGGCTTCTCGCGG 4500  
Db 4441 GTGGGCCAGGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGGCTTCTCGCGG 4500  
QY 4501 GGTACACAGTCTCCCGCAGCATATGTGCTGAGAGGAGCGCTGACAGCATGTCACT 4560  
Db 4501 GGTACACAGTCTCCCGCAGCATATGTGCTGAGAGGAGCGCTGACAGCATGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGGAAGAGGCTTCAACAGTGAATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGGAAGAGGCTTCAACAGTGAATCAAC 4620



[illegible]

Db 6781 ACTAATCTTAAGGGCAGAACTCGCGCTATCGCGGTGCGCGAGCGGTGACTGACG 6840  
QY 6841 ACCAGCTGGGTAAATACCTTCACATGTTACTTGAAGCGCGCTCGCGCTGTCGAGTGG 6900  
Db 6841 ACCAGCTGGGTAAATACCTTCACATGTTACTTGAAGCGCGCTCGCGCTGTCGAGTGG 6900  
QY 6901 AAGCTCCAGACTGACGATGCTGCTATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
Db 6901 AAGCTCCAGACTGACGATGCTGCTATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
QY 6961 GCGGGACCAAGAGCAGCGAGCGACCTACGGGCTTCACGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCAAGAGCAGCGAGCGACCTACGGGCTTCACGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTCGAGTTGATACATCATGCG 7080  
Db 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTCGAGTTGATACATCATGCG 7080  
QY 7081 TCCTCCAATGTGTCAGTCCGCGACGATGCTGCAAAAGGCTGACTATCTCACCCGT 7140  
Db 7081 TCCTCCAATGTGTCAGTCCGCGACGATGCTGCAAAAGGCTGACTATCTCACCCGT 7140  
QY 7141 GACCCACACACCCCTTGGCGGCTGCTGGGAGACGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCACACACCCCTTGGCGGCTGCTGGGAGACGCTAGACACACTCCAGTCAAT 7200  
QY 7201 TCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
Db 7201 TCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTCTCATTCCTCTAGCTCAGAAACACTTGAAGCCCTAGATGTCAG 7320  
Db 7261 ACTCATTTCTCTCATTCCTCTAGCTCAGAAACACTTGAAGCCCTAGATGTCAG 7320  
QY 7321 ATCTAGCGGCTGTTACTCCATTTAGGCACTTACCTACCTCAGATCAATCAACGACTC 7380  
Db 7321 ATCTAGCGGCTGTTACTCCATTTAGGCACTTACCTACCTCAGATCAATCAACGACTC 7380  
QY 7381 CATGGCTTAGCGCATTTTCACTCCATGATTTACTCTCCAGTGAGATCAATAGGCTGCT 7440  
Db 7381 CATGGCTTAGCGCATTTTCACTCCATGATTTACTCTCCAGTGAGATCAATAGGCTGCT 7440  
QY 7441 TCATGCTCAGAAACTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGCGCAGAACT 7500  
Db 7441 TCATGCTCAGAAACTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGCGCAGAACT 7500  
QY 7501 GTCGGCTAGGCTAGCTTCCAGGGGGAGGGCTGCCACTTGTGCCAAGTACCTTCTC 7560  
Db 7501 GTCGGCTAGGCTAGCTTCCAGGGGGAGGGCTGCCACTTGTGCCAAGTACCTTCTC 7560  
QY 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAACTCCGCTGCTGCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAACTCCGCTGCTGCCAGTTGGAT 7620  
QY 7621 TTATCCAGCTGTTGTTGCTGTTTACGCGGGGAGACATATATACAGCCCTGCTCGT 7680  
Db 7621 TTATCCAGCTGTTGTTGCTGTTTACGCGGGGAGACATATATACAGCCCTGCTCGT 7680  
QY 7681 GCCCGACCCGCTGTTGTTGCTGTTTACGCTTACTCTTCTAGGGTAGGCATCTAT 7740  
Db 7681 GCCCGACCCGCTGTTGTTGCTGTTTACGCTTACTCTTCTAGGGTAGGCATCTAT 7740  
QY 7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGGCCATCTGTTTT 7800  
QY 7801 TTTCCCTT 7860  
Db 7801 TTTCCCTT 7860  
QY 7861 TTTTTCCTT 7920  
Db 7861 TTTTTCCTT 7920

QY 7921 TAGCTGTAAAGTCCGTGAGCGCTTGACTGTCAGAGAGTGTGATGCTGGCTCTCTGC 7980  
Db 7921 TAGCTGTAAAGTCCGTGAGCGCTTGACTGTCAGAGAGTGTGATGCTGGCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

RESULT 13  
ABK91435  
ID ABK91435 standard; DNA; 10690 BP.  
XX  
AC ABK91435;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus vector construct pHVNeo.17m2.  
XX  
KW HCV; ss; pHVNeo.17m2; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
XX  
OS Hepatitis C virus.  
OS Encephalomyocarditis virus.  
OS Escherichia coli.  
OS Enterobacteria phage T7.  
OS Synthetic.

Key Location/Qualifiers  
5'UTR 1..341  
CDS 342..1181  
FT misc\_signal /product= "Core-neo fusion protein"  
FT CDS 1190..1800  
FT mutation /tag= a  
FT 3'UTR 759..7991  
FT misc\_feature 7992..10690  
FT WO200259321-A2.  
PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002WO-EP000526.  
XX  
PR 23-JAN-2001; 2001US-0263479P.  
XX  
PA (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
XX  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX  
DR WPI; 2002-599793/64.  
XX  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 16; Page; 69pp; English.

CC The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS4A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV based vector  
 CC pHCVNeo.17 mutant of the invention. Note: The present sequence is not  
 CC shown in the specification but was created by the indexer using the HCV  
 CC vector sequence appearing as ABK91412 and the information in Claim 16  
 XX  
 SQ Sequence 10690 BP; 2334 A; 3044 C; 2908 G; 2404 T; 0 U; 0 Other;

Query Match 100.0%; Score 7985.8; DE 6; Length 10690;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCACGCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60  
 DB 1 GCACGCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60

QY 61 TCTTCACGAGAAAGCGTCTAGCATGGCTTATGATGATGCTGTCGAGCTCCAGGAC 120  
 DB 61 TCTTCACGAGAAAGCGTCTAGCATGGCTTATGATGATGCTGTCGAGCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGATACACCGGAATTGCCAG 180  
 DB 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGATACACCGGAATTGCCAG 180

QY 181 GACGACGGGTCTTTCTTGATCAACCCCTCAATGCCCTGGAGATTGGGGCTGCCCC 240  
 DB 181 GACGACGGGTCTTTCTTGATCAACCCCTCAATGCCCTGGAGATTGGGGCTGCCCC 240

QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGCTACTGCTGATAGG 300  
 DB 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGCTACTGCTGATAGG 300

QY 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCTGACCGTGCACATGAGCAGCAATCCTAAAC 360  
 DB 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCTGACCGTGCACATGAGCAGCAATCCTAAAC 360

QY 361 CTCAGAGAAACCAAGAGGGCGCCCATGATTGAACAGATGGATTGCAGCGTTCTC 420  
 DB 361 CTCAGAGAAACCAAGAGGGCGCCCATGATTGAACAGATGGATTGCAGCGTTCTC 420

QY 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGATCGGCAACAACAGCAATCGGCTGCT 480  
 DB 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGATCGGCAACAACAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGCTGTTCCGGTGTGACGCGAGGGCGCCCGTCTTTTGTCAAGACCG 540  
 DB 481 CTGATGCGCGCTGTTCCGGTGTGACGCGAGGGCGCCCGTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGGTCCCTGATGAATGAATGACAGGAGGAGCGCGGTATCGTGGCTGGCCA 600  
 DB 541 ACCTGTCCGGTCCCTGATGAATGAATGACAGGAGGAGCGCGGTATCGTGGCTGGCCA 600

QY 601 CGACGGGCTTCTTGGCGCAGCTGTGCTCGAGTTGTCACTGAACCGGGAAGGACTGCG 660  
 DB 601 CGACGGGCTTCTTGGCGCAGCTGTGCTCGAGTTGTCACTGAACCGGGAAGGACTGCG 660

QY 661 TGTCTATTGGGCGAAGTGCCTGGGGCAGGATCTCTGTGTCATCTCACCTTCTCTCCGCGA 720  
 DB 661 TGTCTATTGGGCGAAGTGCCTGGGGCAGGATCTCTCTGTGTCATCTCACCTTCTCTCCGCGA 720

QY 721 AAGTATCCATCATGGCTGATGTCAGTGGGGCTGCATAGCTTGTGATCCGGCTACCTGCC 780  
 DB 721 AAGTATCCATCATGGCTGATGTCAGTGGGGCTGCATAGCTTGTGATCCGGCTACCTGCC 780

QY 781 CATTCGACCCCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGAAAGCGGTC 840  
 DB 781 CATTCGACCCCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGAAAGCGGTC 840

QY 841 TTGTGATFAGGATGATCTGCAACGAGAGATCGAGGGCTCGCGCAGCGAAGCTGTCG 900  
 DB 841 TTGTGATFAGGATGATCTGCAACGAGAGATCGAGGGCTCGCGCAGCGAAGCTGTCG 900

QY 901 CAGGCTCAAGGGCGGCTATGCGACGAGAGATCGAGGGCTCGCGCAGCGAAGCTGTCG 960  
 DB 901 CAGGCTCAAGGGCGGCTATGCGACGAGAGATCGAGGGCTCGCGCAGCGAAGCTGTCG 960

QY 961 GCTTCCGGAATATCATGCTGGAAATGCGCGCTTTTCTGGATTCATCGACTGTGCGCGC 1020  
 DB 961 GCTTCCGGAATATCATGCTGGAAATGCGCGCTTTTCTGGATTCATCGACTGTGCGCGC 1020

QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATTTGCTGAAGAC 1080  
 DB 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATTTGCTGAAGAC 1080

QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTCCCGATTCGC 1140  
 DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTCCCGATTCGC 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTCTTCTGAGTTTAAACAGACCAACG 1200  
 DB 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTCTTCTGAGTTTAAACAGACCAACG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTACCTTACTGCG 1260  
 DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTACCTTACTGCG 1260

QY 1261 CGAAGCGCTTGAATTAAGCGCGGTGTGCGTCTATATGTTATTTTCCACCATATG 1320  
 DB 1261 CGAAGCGCTTGAATTAAGCGCGGTGTGCGTCTATATGTTATTTTCCACCATATG 1320

QY 1321 CCGTCTTTTGGCAATGTAGGGCGCGGAAACCTTGGCCCTCTCTTCTGACGAGCATTCCT 1380  
 DB 1321 CCGTCTTTTGGCAATGTAGGGCGCGGAAACCTTGGCCCTCTCTTCTGACGAGCATTCCT 1380

QY 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGATGTGCTGTAAGTAAAGCA 1440  
 DB 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGATGTGCTGTAAGTAAAGCA 1440

QY 1441 GTTTCCTCTGAAGTCTTTTGAAGCAACAAACGCTGTAGCGACCTTTTGCAGGACGG 1500  
 DB 1441 GTTTCCTCTGAAGTCTTTTGAAGCAACAAACGCTGTAGCGACCTTTTGCAGGACGG 1500

QY 1501 AACCCCGACCTGCGGACAGGTGCTCTGCGGCGCAAGCAACGCTGTATAGTACACCT 1560  
 DB 1501 AACCCCGACCTGCGGACAGGTGCTCTGCGGCGCAAGCAACGCTGTATAGTACACCT 1560

QY 1561 GCAAGGGCGGCAAAACCCAGTGCACGTTGTGAGTTGATGTGGAAGAGTCAAA 1620  
 DB 1561 GCAAGGGCGGCAAAACCCAGTGCACGTTGTGAGTTGATGTGGAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCGTATTCAAAGGGCTGAAGGATGCCAGAGGTATCCCATTTGT 1680  
 DB 1621 TGGCTCTCTCAAGCGTATTCAAAGGGCTGAAGGATGCCAGAGGTATCCCATTTGT 1680

QY 1681 ATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTGTTTGTGAGGTTAA 1740  
 DB 1681 ATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGTGTTTGTGAGGTTAA 1740

QY 1741 AACGTCTAGGCCCCCGGAAACCGAGGGAGCGTGGTTTCTTTGAAACACGATAATACC 1800

Db 1741 AAGCTCTAGGCCCCCGAACCACGAGGAGCTGGTTTCTCTTCAAAAAACAGATATACC 1800  
QY 1801 ATGGGCGCTATTACGGCTTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGGCGCTATTACGGCTTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCAAGGCGGGACAGGAACACAGGTGAGGGGGAGGTCAAGTGTCTTCCACCGCA 1920  
Db 1861 AGCCTCAAGGCGGGACAGGAACACAGGTGAGGGGGAGGTCAAGTGTCTTCCACCGCA 1920  
QY 1921 ACACAACTCTTCTGGGACCTGCTGCTCAATGCGGTGTGTGGACTGTCTATCATGTGCC 1980  
Db 1921 ACACAACTCTTCTGGGACCTGCTGCTCAATGCGGTGTGTGGACTGTCTATCATGTGCC 1980  
QY 1981 GGCTCAAAAGACCTTGGCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGAC 2040  
Db 1981 GGCTCAAAAGACCTTGGCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGAC 2040  
QY 2041 CAGGACCTCTGGGTGCGCAAGCGCCCCCGGGGCGGTTCCTTGACACCAATGACACCTGC 2100  
Db 2041 CAGGACCTCTGGGTGCGCAAGCGCCCCCGGGGCGGTTCCTTGACACCAATGACACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTACTTGGTCAAGAGGATGCGCGATGCAATTCGGGTGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTACTTGGTCAAGAGGATGCGCGATGCAATTCGGGTGCGCGCGG 2160  
QY 2161 GGCACAGCAGGGGAGGCTACTCTCCCCCAGGCGCGTCTCTACTTGAAGGGGTCTTTCG 2220  
Db 2161 GGCACAGCAGGGGAGGCTACTCTCCCCCAGGCGCGTCTCTACTTGAAGGGGTCTTTCG 2220  
QY 2221 GGGGCTGCACTGCTGTCGCCCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGTGTCG 2280  
Db 2221 GGGGCTGCACTGCTGTCGCCCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGTGTCG 2280  
QY 2281 ACCGAGGGGTGCGAAGGGGTGGAATTTGTACCGTTCGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTGCGAAGGGGTGGAATTTGTACCGTTCGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGCGGCTAAGGTGCGGCTGGGTATGCA 2460  
Db 2341 CGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGCGGCTAAGGTGCGGCTGGGTATGCA 2460  
QY 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCA 2460  
Db 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCA 2460  
QY 2461 GCCCAAGGGTATAGGTGCTTGTCTGAAACCGGTTCGGTCCGCGCCACCTTAGTTTCGGG 2520  
Db 2461 GCCCAAGGGTATAGGTGCTTGTCTGAAACCGGTTCGGTCCGCGCCACCTTAGTTTCGGG 2520  
QY 2521 GGGTATATGCTTAAGGCACATGTATCGACCTAACATCAGAACCGGGGTAAAGCACTC 2580  
Db 2521 GGGTATATGCTTAAGGCACATGTATCGACCTAACATCAGAACCGGGGTAAAGCACTC 2580  
QY 2581 ACACGGGTGCCCTTACGACTTCCACTATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2581 ACACGGGTGCCCTTACGACTTCCACTATGAGTGCACCTCAACTGACTCGACCACT 2700  
QY 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
QY 2701 ATCTTGGGCGCTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2760  
Db 2701 ATCTTGGGCGCTATGACATCATATATGATGAGTGCACCTCAACTGACTCGACCACT 2760  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTACCGTGCCACATCCAAACATCGAGAGGTG 2820  
Db 2761 CTCGCCACCGCTACGCTCCGGGATCGGTACCGTGCCACATCCAAACATCGAGAGGTG 2820  
QY 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCACTC 2880  
Db 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCACTC 2880

Db 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCACTC 2880  
QY 2881 AAGGGGGGAGGACACTCATTTTCTGCCATTCCAAGAAAGAAATGTATGAGTTCGGCGG 2940  
Db 2881 AAGGGGGGAGGACACTCATTTTCTGCCATTCCAAGAAAGAAATGTATGAGTTCGGCGG 2940  
QY 2941 AAGCTGTCCGGCTCGACTCAATCTGTAGCATATTAACGGGGGCTTGTATGTATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGACTCAATCTGTAGCATATTAACGGGGGCTTGTATGTATCCGTC 3000  
QY 3001 ATACCAACTAGCGGAGAGCTCATTTGTGTAGCAACGAGCGCTCTAATGACGGCTTTTACC 3060  
Db 3001 ATACCAACTAGCGGAGAGCTCATTTGTGTAGCAACGAGCGCTCTAATGACGGCTTTTACC 3060  
QY 3061 GGCATTTTCCGACTCAGTGTGATCGACTGCAATCATGTGTCAACGAGCGCTCTAATGACGGCTTTTACC 3120  
Db 3061 GGCATTTTCCGACTCAGTGTGATCGACTGCAATCATGTGTCAACGAGCGCTCTAATGACGGCTTTTACC 3120  
QY 3121 CTGACCCGACCTTCACTTCACTTGAAGACGACGCGGTGCGCAAGACGCGGTGTCAAGCTGC 3180  
Db 3121 CTGACCCGACCTTCACTTCACTTGAAGACGACGCGGTGCGCAAGACGCGGTGTCAAGCTGC 3180  
QY 3181 CAGCGGCGAGGACGAGCTGTGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGGCGAGGACGAGCTGTGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAAACGGGCTCGGGCATGTTCCGATTCCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Db 3241 GAAACGGGCTCGGGCATGTTCCGATTCCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
QY 3301 GCTTGTGACGAGCTCACGCCCGCGAGACCTCACTAGTTAGTTGCGGGCTTACTTAAACACA 3360  
Db 3301 GCTTGTGACGAGCTCACGCCCGCGAGACCTCACTAGTTAGTTGCGGGCTTACTTAAACACA 3360  
QY 3361 CCAAGGTTGCGGCTCTGCGAGGACCATCTGAGGTTCTGGAGAGCGCTTTTACAGGCGCTC 3420  
Db 3361 CCAAGGTTGCGGCTCTGCGAGGACCATCTGAGGTTCTGGAGAGCGCTTTTACAGGCGCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGGACACACTTCCCTTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGGACACACTTCCCTTAC 3480  
QY 3481 CTGGTACATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
Db 3481 CTGGTACATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
QY 3541 CAAATGTGAGTGTCTCATCGGCTAAAGCTTACGCTGCAAGGCGCAACGCCCTTGTG 3600  
Db 3541 CAAATGTGAGTGTCTCATCGGCTAAAGCTTACGCTGCAAGGCGCAACGCCCTTGTG 3600  
QY 3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTTACTACCAACACCCCATACCAATACATC 3660  
Db 3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTTACTACCAACACCCCATACCAATACATC 3660  
QY 3661 ATGGCATGATCGGCTGACCTGAGGTCGTCAAGGACCTGGTGTGTTAGGCGGA 3720  
Db 3661 ATGGCATGATCGGCTGACCTGAGGTCGTCAAGGACCTGGTGTGTTAGGCGGA 3720  
QY 3721 GTCTTAGCAGCTCTGGCGCGTATTGCTGACAAAGGCGTGGTCAATGTGAGGCGAG 3780  
Db 3721 GTCTTAGCAGCTCTGGCGCGTATTGCTGACAAAGGCGTGGTCAATGTGAGGCGAG 3780  
QY 3781 ATCATCTTGTTCGGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTTCGGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGTGGAGAGTGGCGCTTCAACCTCTTACATCGAACAGGGAATTCAGCTCGCC 3900  
Db 3841 GATGAGTGGAGAGTGGCGCTTCAACCTCTTACATCGAACAGGGAATTCAGCTCGCC 3900  
QY 3901 GAACTTTCAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
Db 3901 GAACTTTCAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960



Db	6121	CTGCGCGCAGAGAGAGGTCACTTTGACAGACTGCAGGTCTCTGACGACCACTACCGGAC	6180
QY	6181	GTGCTCAAGGAGATGAAGCGCAAGCGCTCCACAGTTAAGGCTAACTTCATCCGTGGAG	6240
Db	6181	GTGCTCAAGGAGATGAAGCGCAGAGCGCTCCACAGTTAAGGCTAACTTCATCCGTGGAG	6240
QY	6241	GNAGCCTGTAGCTGACAGCGCCGCCACATNTGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Db	6241	GAAGCCTGTAGCTGACGCGCCGCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
QY	6301	GACGTCGGGAACCTATCTCAGCAAGCGCGTTAACACATTCGCTCCGTGTGGAAAGGACTTG	6360
Db	6301	GACGTCGGGAACCTATCTCAGCAAGCGCGTTAACACATTCGCTCCGTGTGGAAAGGACTTG	6360
QY	6361	CTGAAGACACTGTAGACACCAATTTGACACCACTCATGCGCAAAATGAGGTTTCTGC	6420
Db	6361	CTGAAGACACTGTAGACACCAATTTGACACCACTCATGCGCAAAATGAGGTTTCTGC	6420
QY	6421	GTCCAAACAGAAAGGGGGCGCGCAAGCTCGCCTTATCGCTATTCGATTTCCAGATTTGGGG	6480
Db	6421	GTCCAAACAGAAAGGGGGCGCGCAAGCTCGCCTTATCGCTATTCGATTTCCAGATTTGGGG	6480
QY	6481	GTTCTGTGTGCGAGAAATAGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG	6540
Db	6481	GTTCTGTGTGCGAGAAATAGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG	6540
QY	6541	ATGGGCTCTTCATACGGATTTCCAATCTCTCTCTGGACAGCGGTTCGAGTTCTCTGGTGAAT	6600
Db	6541	ATGGGCTCTTCATACGGATTTCCAATCTCTCTCTGGACAGCGGTTCGAGTTCTCTGGTGAAT	6600
QY	6601	GCCTGGAAGGGAAGAAATGCCCTATGGGCTTCGCATATGACACCGCGCTTTTGACTCA	6660
Db	6601	GCCTGGAAGGGAAGAAATGCCCTATGGGCTTCGCATATGACACCGCGCTTTTGACTCA	6660
QY	6661	ACGCTCACTGAGATGACATCCGTGTGAGGACTCAATCTACCATTTGTGTGACTTGGCC	6720
Db	6661	ACGCTCACTGAGATGACATCCGTGTGAGGACTCAATCTACCATTTGTGTGACTTGGCC	6720
QY	6721	CCGGAAGCCAGACAGGCCATAAAGTCCGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
Db	6721	CCGGAAGCCAGACAGGCCATAAAGTCCGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
QY	6781	ACTAATTTCTAAAGGCGAGAACTCGCGGTATCGCGCGTCCGCGAGCGGTGACTGACG	6840
Db	6781	ACTAATTTCTAAAGGCGAGAACTCGCGGTATCGCGCGTCCGCGAGCGGTGACTGACG	6840
QY	6841	ACCAGCTGGCGTAAATACCCCTCACATGTACTTTGAAGCGCGCTGCGGCGCTGCGAGCTCG	6900
Db	6841	ACCAGCTGGCGTAAATACCCCTCACATGTACTTTGAAGCGCGCTGCGGCGCTGCGAGCTCG	6900
QY	6901	AAGCTCAGACTGACAGATGCTCGTATCGGAGACGACTTGTCTGTTATCTGTGAAAGC	6960
Db	6901	AAGCTCAGACTGACAGATGCTCGTATCGGAGACGACTTGTCTGTTATCTGTGAAAGC	6960
QY	6961	GGGGGACCCAGACAGACGAGCGAGCTTACGGGCTTACGGAGGCTATGACTAGATAC	7020
Db	6961	GGGGGACCCAGACAGACGAGCGAGCTTACGGGCTTACGGAGGCTATGACTAGATAC	7020
QY	7021	TCTGCCCCCTTGGGACCCGCCCAACAGAAATACGACTTGGAGTTGATTAACATATGC	7080
Db	7021	TCTGCCCCCTTGGGACCCGCCCAACAGAAATACGACTTGGAGTTGATTAACATATGC	7080
QY	7081	TCTTCCAATGTGTAGTCCGACGATGCATCTGCGCAAAAGGTTGATCTCTCACCCGT	7140
Db	7081	TCTTCCAATGTGTAGTCCGACGATGCATCTGCGCAAAAGGTTGATCTCTCACCCGT	7140
QY	7141	GACCCACACCCCTTGGCGGGCTGCTGGAGACAGCTAGACACACTCCAGTCAAT	7200
Db	7141	GACCCACACCCCTTGGCGGGCTGCTGGAGACAGCTAGACACACTCCAGTCAAT	7200
QY	7201	TCTGGCTAGGCAACATCATATGATGCGCCCACTTGTGGGCAAGGATGATCCTGATG	7260

Db	7201	TCCTGGCTAGGCAACATCATGTATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
QY	7261	ACTCATTTCTTCATCCTCTAGCTCAGAAACAATTGAAAAGCCCTAGATTGTCAAG	7320
Db	7261	ACTCATTTCTTCATCCTCTAGCTCAGAAACAATTGAAAAGCCCTAGATTGTCAAG	7320
QY	7321	AFCTACGGGGCGTGTACTCCATTGAGCCACTTGACCTTACCTCAGATCAATTCACGACTC	7380
Db	7321	ATCTACGGGGCGTGTACTCCATTGAGCCACTTGACCTCAGATCAATTCACGACTC	7380
QY	7381	CATGGCCCTTAGGGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
Db	7381	CATGGCCCTTAGGGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
QY	7441	TCATGCCCTCAGAAACTTTGGGTTACCGGCCCTTGCAGTCTGGAGACATCGGGCCAGAAGT	7500
Db	7441	TCATGCCCTCAGAAACTTTGGGTTACCGGCCCTTGCAGTCTGGAGACATCGGGCCAGAAGT	7500
QY	7501	GTCGCGCTAGCGTACTGTCCTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
Db	7501	GTCGCGCTAGCGTACTGTCCTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
QY	7561	AACTGGGCAGTAAGAGCAAAAGCTCAAACTCACATCCCAATCCCGCTGGCTCCCAAGTTGGAT	7620
Db	7561	AACTGGGCAGTAAGAGCAAAAGCTCAAACTCACATCCCAATCCCGCTGGCTCCCAAGTTGGAT	7620
QY	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCCCTGCTCGT	7680
Db	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCCCTGCTCGT	7680
QY	7681	GCCCGACCCCGCTGGTTCAATGTGTGCTACTCCTACTTCTGTAGGGGTAGGCATCTAT	7740
Db	7681	GCCCGACCCCGCTGGTTCAATGTGTGCTACTCCTACTTCTGTAGGGGTAGGCATCTAT	7740
QY	7741	CTACTCCCGAACGATGAACGGGGACCTAAACACTCCAGGCCAAATAGGCCATCTGTTTT	7800
Db	7741	CTACTCCCGAACGATGAACGGGGACCTAAACACTCCAGGCCAAATAGGCCATCTGTTTT	7800
QY	7801	TTTTCCCTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCTTTT	7860
Db	7801	TTTTCCCTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCTTTT	7860
QY	7861	TTTTTCTCTTTTTTCTTTTTTCTTTCCCTTTTCTTTCCCTTTCGTTGGCTCCATCTTAGCCCTAGTCAACGGC	7920
Db	7861	TTTTTCTCTTTTTTCTTTTTTCTTTCCCTTTTCTTTCCCTTTCGTTGGCTCCATCTTAGCCCTAGTCAACGGC	7920
QY	7921	TAGCTGAAAGTCCGTTAGCGCGCTTGACTCAGAGAGTGTCTGATACCTGGCTCTCTGCG	7980
Db	7921	TAGCTGAAAGTCCGTTAGCGCGCTTGACTCAGAGAGTGTCTGATACCTGGCTCTCTGCG	7980
QY	7981	AGATCAAGT	7989
Db	7981	AGATCAAGT	7989

RESULT 14	
ABK91243	
ID	ABK91243 standard; DNA; 10690 BP.
XX	
XX	
XX	ABK91243;
DT	15-NOV-2002 (first entry)
XX	
DE	Hepatitis C virus vector construct
XX	
KW	HCV; ss; pHCVNeo.17m2; adaptive m
KW	hepatocellular carcinoma; NS3; NS5
KW	internal ribosome entry site; IRE
XX	
XX	Hepatitis C virus.
OS	Encephalomyocarditis virus.
OS	Escherichia coli.
OS	Enterobacteria phage T7.







QY 1081 TTGGGGCGGAATGGGTGACCGCTTCTCGTGTGTTTACGGTATCGCGCTCCGATTTCG 1140  
Db 1081 TTGGGGCGGAATGGGTGACCGCTTCTCGTGTGTTTACGGTATCGCGCTCCGATTTCG 1140  
QY 1141 AGGGATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG 1200  
Db 1141 AGGGATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCGCTCCCGCCCGCTTAACGTTACTGCG 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCGCTCCCGCCCGCTTAACGTTACTGCG 1260  
QY 1261 CGAAGCGCTTGAATTAAGCGCGTGTGCGTTTGTCTATATGTTTTCACCAATATG 1320  
Db 1261 CGAAGCGCTTGAATTAAGCGCGTGTGCGTTTGTCTATATGTTTTCACCAATATG 1320  
QY 1321 CGCTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTGACGAGCATTCCT 1380  
Db 1321 CGCTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTGACGAGCATTCCT 1380  
QY 1381 AGGGTCTTTCCCTCTCGCAAGGAATGCAAGTCTGTTGAATGTCTGTAAGTGTCT 1440  
Db 1381 AGGGTCTTTCCCTCTCGCAAGGAATGCAAGTCTGTTGAATGTCTGTAAGTGTCT 1440  
QY 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACACGTCGTAGCGACCTTTTGACGCGCG 1500  
Db 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACACGTCGTAGCGACCTTTTGACGCGCG 1500  
QY 1501 AACCCCGACCTGGGACAGGTGCTCTGGGCGCAAAAGCCACGTTGATAACACCT 1560  
Db 1501 AACCCCGACCTGGGACAGGTGCTCTGGGCGCAAAAGCCACGTTGATAACACCT 1560  
QY 1561 GCAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGAATGTGTAAGTGTCTGAA 1620  
Db 1561 GCAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGAATGTGTAAGTGTCTGAA 1620  
QY 1621 TGGCTCTCTCAAGCTATTCAACAGGGGCTGGAAGGATGCCAGAGTACCCATTGT 1680  
Db 1621 TGGCTCTCTCAAGCTATTCAACAGGGGCTGGAAGGATGCCAGAGTACCCATTGT 1680  
QY 1681 ATGGATCTGATCTGGGCTCTGGTGACATGCTTTTACATGTTTGTAGTTCGAGTTAAA 1740  
Db 1681 ATGGATCTGATCTGGGCTCTGGTGACATGCTTTTACATGTTTGTAGTTCGAGTTAAA 1740  
QY 1741 AACCTCTAGGCCCCCGAAACACGGGAGCGTGTTCCTTTGAAACACGATATACC 1800  
Db 1741 AACCTCTAGGCCCCCGAAACACGGGAGCGTGTTCCTTTGAAACACGATATACC 1800  
QY 1801 ATGGCGCTATTAGCGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGCGCTATTAGCGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCTCACAGGCGGGACAGAACCGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTCACAGGCGGGACAGAACCGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACAATCTTCTGGGACCTGGTCAATGGGCTGTGGTCTGCTATCATGTGTC 1980  
Db 1921 ACACAATCTTCTGGGACCTGGTCAATGGGCTGTGGTCTGCTATCATGTGTC 1980  
QY 1981 GGCTCAAGACCTTTCGCGCCCAAGGGGCCAATACACCAAAATGTACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCTTTCGCGCCCAAGGGGCCAATACACCAAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACCTCTGGCTGGGAAGCGCCCGCGGCGGCTTCTTGACACCAATGCATGTC 2100  
Db 2041 CAGGACCTCTGGCTGGGAAGCGCCCGCGGCGGCTTCTTGACACCAATGCATGTC 2100  
QY 2101 GGCAGTCTGGACCTTTACTTGTGTCAGAGCATGCGCATGTCTTCGGTGGCGCGG 2160  
Db 2101 GGCAGTCTGGACCTTTACTTGTGTCAGAGCATGCGCATGTCTTCGGTGGCGCGG 2160  
QY 2161 GGCACAGCAGGGGAGCCTACTCTCCCGAGCGCGCTCTCTTACTTGAAGGCTCTTCG 2220

Db 2161 GGCACAGCAGGGGAGCCTACTCTCCCGAGCGCGCTCTCTTACTTGAAGGCTCTTCG 2220  
QY 2221 GCGGTCTCACTCTCTGCCCCCTCGGGGACGCTGTGGGATCTTTTCGGCTGCGGTGC 2280  
Db 2221 GCGGTCTCACTCTCTGCCCCCTCGGGGACGCTGTGGGATCTTTTCGGCTGCGGTGC 2280  
QY 2281 ACCGAGGGGTTCGAAAGCGGTGGAATTTGPAACCGCTGAGTCTATGTAACCACTATG 2340  
Db 2281 ACCGAGGGGTTCGAAAGCGGTGGAATTTGPAACCGCTGAGTCTATGTAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTTCACGACAACTCTGCTCCCTCGGCGGTACCGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTTCACGACAACTCTGCTCCCTCGGCGGTACCGACATTCAGGTG 2400  
QY 2401 GCCATCTACACGCGCTTACTGTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
Db 2401 GCCATCTACACGCGCTTACTGTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
QY 2461 GCCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGTCCGCGCACCTAGGTTCGGG 2520  
Db 2461 GCCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGTCCGCGCACCTAGGTTCGGG 2520  
QY 2521 GCGTATATGTTAAGGACATGTTATCGACCTTAACATCAGAACCGGGGTAGGACCATC 2580  
Db 2521 GCGTATATGTTAAGGACATGTTATCGACCTTAACATCAGAACCGGGGTAGGACCATC 2580  
QY 2581 ACCACGGGTGCCCCCTACCTACTCAGCTACTCCACCTATGCGAAGTTCCTGCGGCGGTGTC 2640  
Db 2581 ACCACGGGTGCCCCCTACCTACTCAGCTACTCCACCTATGCGAAGTTCCTGCGGCGGTGTC 2640  
QY 2641 TCTGGGCGGCTTATGATCATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGCGGCTTATGATCATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCTTCGACACAGTCTCGGACCAAGCGGAGACGCTGAGCGCGACTCGTGTG 2760  
Db 2701 ATCTGGGCTTCGACACAGTCTCGGACCAAGCGGAGACGCTGAGCGCGACTCGTGTG 2760  
QY 2761 CTGCGCACCGCTAGCGCTCGGGATCGGTACCGTGCACATCCCAATCGAGACCATC 2820  
Db 2761 CTGCGCACCGCTAGCGCTCGGGATCGGTACCGTGCACATCCCAATCGAGAGGTG 2820  
QY 2821 GCTCTGTCCAGCATGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCATGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGAGGACCTCAATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG 2940  
Db 2881 AAGGGGGAGGACCTCAATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG 2940  
QY 2941 AAGCTGTCCGCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATTCGTC 3000  
Db 2941 AAGCTGTCCGCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATTCGTC 3000  
QY 3001 ATACCAATAGCGGACGCTCAATTTCTGAGCAACGCACTCTAATGACGGCTTTACC 3060  
Db 3001 ATACCAATAGCGGACGCTCAATTTCTGAGCAACGCACTCTAATGACGGCTTTACC 3060  
QY 3061 GCGATTTTCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
Db 3061 GCGATTTTCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120  
QY 3121 CTGACCGGACCTTACCATTTGAGACGACGCTGCGGACGAGCGGTGTCACGCTCG 3180  
Db 3121 CTGACCGGACCTTACCATTTGAGACGACGCTGCGGACGAGCGGTGTCACGCTCG 3180  
QY 3181 CAGCGGAGGAGGACTGTTAGGGGACGATGGGCAATTTACAGGTTTGTGCTCCAGGA 3240  
Db 3181 CAGCGGAGGAGGACTGTTAGGGGACGATGGGCAATTTACAGGTTTGTGCTCCAGGA 3240  
QY 3241 GAAAGGCGCTCGGGCATGTTCTCGTCTGTGCGAGTGTATGACGCGGCTGT 3300

Db 3241 GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGGAGTGCTATGACGCGGGCTGT 3300  
Qy  
3301 GCTTGGTACGAGCTCACGCCCGCGGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
Db  
3302 GCTTGGTACGAGCTCACGCCCGCGGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
Qy  
3361 CGAGGGTTCGGCTGTGCGAGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGCTC 3420  
Db  
3362 CGAGGGTTCGGCTGTGCGAGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGCTC 3420  
Qy  
3421 ACCACATAGAGCCCATTTCTGTCCCAAGCTAAGACGCGGAGAGACAACTTCCCTAC 3480  
Db  
3422 ACCACATAGAGCCCATTTCTGTCCCAAGCTAAGACGCGGAGAGACAACTTCCCTAC 3480  
Qy  
3481 CTGTGTAGCATACAGGCTACGTTGTGCGGACAGGCTCAGGCTCCACCTCATCGTGGAC 3540  
Db  
3482 CTGTGTAGCATACAGGCTACGTTGTGCGGACAGGCTCAGGCTCCACCTCATCGTGGAC 3540  
Qy  
3541 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTACGCTGACGGGCCAACGCCCTGCTG 3600  
Db  
3542 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTACGCTGACGGGCCAACGCCCTGCTG 3600  
Qy  
3601 TATAGGCTGGAGCGCTTCAAAACGAGGTTACTTACACACACCCCAATAACCAATACATC 3660  
Db  
3602 TATAGGCTGGAGCGCTTCAAAACGAGGTTACTTACACACACCCCAATAACCAATACATC 3660  
Qy  
3661 ATGGCATGATGTGCGCTGACCTGGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720  
Db  
3662 ATGGCATGATGTGCGCTGACCTGGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720  
Qy  
3721 GTCTTAGAGCTGTGCGCGGTATTCCTGACAAACAGGAGCGTGTCTATTGGGCGAGG 3780  
Db  
3722 GTCTTAGAGCTGTGCGCGGTATTCCTGACAAACAGGAGCGTGTCTATTGGGCGAGG 3780  
Qy  
3781 ATCATCTTGTCCGGAAGCGGCGCATCAITCCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db  
3782 ATCATCTTGTCCGGAAGCGGCGCATCAITCCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Qy  
3841 GATGAGATGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCGAGCTGCC 3900  
Db  
3901 GAACAATTCACACAGAGGAATCGGGTCTGCAACAGCCACCAAGCAAGCGGAGGCT 3960  
Qy  
3902 GAACAATTCACACAGAGGAATCGGGTCTGCAACAGCCACCAAGCAAGCGGAGGCT 3960  
Qy  
3961 GCTGCTCCCTGTGTGAATCCAAAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db  
3962 GCTGCTCCCTGTGTGAATCCAAAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Qy  
4021 TGGAAATTCATACGCGGATACAAATTTTAGAGGCTTTGCTCACTCTGCTGCGCAACCCC 4080  
Db  
4022 TGGAAATTCATACGCGGATACAAATTTTAGAGGCTTTGCTCACTCTGCTGCGCAACCCC 4080  
Qy  
4081 GCGATAGCATCACTGATGGCAATTCACAGCCTTATATCACAGCCCGCTCACCAACCAAT 4140  
Db  
4082 GCGATAGCATCACTGATGGCAATTCACAGCCTTATATCACAGCCCGCTCACCAACCAAT 4140  
Qy  
4141 ACCCTCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
Db  
4142 ACCCTCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
Qy  
4201 GCTTCTGTTTCTGAGGCGCGGATCGCTGGAGCGGCTGTGCGAGATAGGCTTGGG 4260  
Db  
4202 GCTTCTGTTTCTGAGGCGCGGATCGCTGGAGCGGCTGTGCGAGATAGGCTTGGG 4260  
Qy  
4261 AAGGTGCTTGGGATATTTTGGAGGTTATGAGCAGGGGTGGAGCGGCGCTCGTGGCC 4320  
Db  
4262 AAGGTGCTTGGGATATTTTGGAGGTTATGAGCAGGGGTGGAGCGGCGCTCGTGGCC 4320  
Qy  
4321 TTTAAGGTATGAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAAGCTACTCCCTGCT 4380  
Db  
4322 TTTAAGGTATGAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAAGCTACTCCCTGCT 4380

Qy 4381 ATCTCTCCCTCGGCGCCTAGTCTCGGGGTCTGTGTCGACGACATCTGCTGCGGCAC 4440  
Db  
4382 ATCTCTCCCTCGGCGCCTAGTCTCGGGGTCTGTGTCGACGACATCTGCTGCGGCAC 4440  
Qy  
4441 GTGGGCCACAGGGAGGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCGCTTCGCGG 4500  
Db  
4442 GTGGGCCACAGGGAGGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCGCTTCGCGG 4500  
Qy  
4501 GGTAAACAGCTCTCCCCACGACATATGTGCTGAGAGCGACGCTGACAGCGTGTCACT 4560  
Db  
4502 GGTAAACAGCTCTCCCCACGACATATGTGCTGAGAGCGACGCTGACAGCGTGTCACT 4560  
Qy  
4561 CAGATCTCTCTPAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
Db  
4562 CAGATCTCTCTPAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
Qy  
4621 GAGGACTGTCCACGCCATGTCTCCGCTCTGAGAGTGAAGAGATGTTTGGGATGGATATGC 4680  
Db  
4622 GAGGACTGTCCACGCCATGTCTCCGCTCTGAGAGTGAAGAGATGTTTGGGATGGATATGC 4680  
Qy  
4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCTGCGCGATTTGCCGGA 4740  
Db  
4682 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCTGCGCGATTTGCCGGA 4740  
Qy  
4741 GTCCCCCTTCTCTCATGTCAACGTGGGTAAAGGAGTCTGGCGGGCGACGSCATCATG 4800  
Db  
4742 GTCCCCCTTCTCTCATGTCAACGTGGGTAAAGGAGTCTGGCGGGCGACGSCATCATG 4800  
Qy  
4801 CAAACACCTGCCATGTGGAGCAGATCACCGGACATGTGAAACACCGGTTCCATGAGG 4860  
Db  
4802 CAAACACCTGCCATGTGGAGCAGATCACCGGACATGTGAAACACCGGTTCCATGAGG 4860  
Qy  
4861 ATGCTGGGCGCTAGGACCTGTAGTAAACGTGTCATGGAACATTTCCCAATTAACGCGTAC 4920  
Db  
4862 ATGCTGGGCGCTAGGACCTGTAGTAAACGTGTCATGGAACATTTCCCAATTAACGCGTAC 4920  
Qy  
4921 ACCACGGGCGCTTGCACGCGCTTCCCGGCGCCAAATTTATTTAGGGGCTGTGGCGGCTG 4980  
Db  
4922 ACCACGGGCGCTTGCACGCGCTTCCCGGCGCCAAATTTATTTAGGGGCTGTGGCGGCTG 4980  
Qy  
4981 GCTGCTGAGAGTACGTTGAGGTTTACGCGGTGGGGATTTCCACTACGTGACGGGCTG 5040  
Db  
4982 GCTGCTGAGAGTACGTTGAGGTTTACGCGGTGGGGATTTCCACTACGTGACGGGCTG 5040  
Qy  
5041 ACCACTGACAACTGAAGTCCCGTGTACGTTCCGCGCCCGAAATTTCTTTCACAGAAAGT 5100  
Db  
5042 ACCACTGACAACTGAAGTCCCGTGTACGTTCCGCGCCCGAAATTTCTTTCACAGAAAGT 5100  
Qy  
5101 GATGGGCTGCGGTTGCACAGGTACGTTCCAGCGTGAACACCCCTCTTACCGGAGGAGTTC 5160  
Db  
5102 GATGGGCTGCGGTTGCACAGGTACGTTCCAGCGTGAACACCCCTCTTACCGGAGGAGTTC 5160  
Qy  
5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGAGCCCGAA 5220  
Db  
5162 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGAGCCCGAA 5220  
Qy  
5221 CCGGACGTAGAGTGTCACTTCCATGCTCAACGACCCCTCCCAATTAAGCGGAGAGG 5280  
Db  
5222 CCGGACGTAGAGTGTCACTTCCATGCTCAACGACCCCTCCCAATTAAGCGGAGAGG 5280  
Qy  
5281 GCTAAGCGTAGGCTGGCGAGGATCTCCCGCTCTTGGCCAGCTCATGAGTGGAG 5340  
Db  
5282 GCTAAGCGTAGGCTGGCGAGGATCTCCCGCTCTTGGCCAGCTCATGAGTGGAG 5340  
Qy  
5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCGTGCAC 5400  
Db  
5342 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCGTGCAC 5400  
Qy  
5401 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCGTGCAC 5460  
Db  
5402 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCGTGCAC 5460





QY 421 CGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCGACAAACAGACAATTCGGCTGCT 480  
Db 421 CGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCGACAAACAGACAATTCGGCTGCT 480  
QY 481 CTGATGCGCGCTGTTCGGCTGTTCAGCGAGGGGCGCCGGTCTTTTGTTCGAAGACCG 540  
Db 481 CTGATGCGCGCTGTTCGGCTGTTCAGCGAGGGGCGCCGGTCTTTTGTTCGAAGACCG 540  
QY 541 ACCTGTCCGGTCCCTGAATGAATGCACTGAGGAGGAGGAGCGCGGCTATTCGTGGTGGCCA 600  
Db 541 ACCTGTCCGGTCCCTGAATGAATGCACTGAGGAGGAGGAGCGCGGCTATTCGTGGTGGCCA 600  
QY 601 CGACGGGCTTCCTTGGCGACGTGTCTCGACGTGTTCACATGAAGCGGAGGAGGAGTGGC 660  
Db 601 CGACGGGCTTCCTTGGCGACGTGTCTCGACGTGTTCACATGAAGCGGAGGAGGAGTGGC 660  
QY 661 TGTATTTGGGCGAAGTGCAGGCGAGGATCTCTGTGTATCTACCTTGTCTCTCTCCCGAGA 720  
Db 661 TGTATTTGGGCGAAGTGCAGGCGAGGATCTCTGTGTATCTACCTTGTCTCTCTCCCGAGA 720  
QY 721 AAGTATCCATCATCGCTGATGCAATGCGCGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
Db 721 AAGTATCCATCATCGCTGATGCAATGCGCGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
QY 781 CATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGCGTC 840  
Db 781 CATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGCGTC 840  
QY 841 TTGTGCGATCAGATGATCTCGAAGAGCATCAGGGGCTCGGCGCACCGCAACTGTTCG 900  
Db 841 TTGTGCGATCAGATGATCTCGAAGAGCATCAGGGGCTCGGCGCACCGCAACTGTTCG 900  
QY 901 CCAGGCTCAAGGCGCGATCGCGAGCGGAGGATCTCTGCTGAGCCATGGCGATGCCCT 960  
Db 901 CCAGGCTCAAGGCGCGATCGCGAGCGGAGGATCTCTGCTGAGCCATGGCGATGCCCT 960  
QY 961 GCTTGGCGAATATCATGGTGGAAATGGCGCTTTCTGATTCATCGACTGTGGCGGC 1020  
Db 961 GCTTGGCGAATATCATGGTGGAAATGGCGCTTTCTGATTCATCGACTGTGGCGGC 1020  
QY 1021 TGGGTGGCGGACCGGTATCAGGATAGCTGTGGCTACCCGTGATTTGCTGAAGAGC 1080  
Db 1021 TGGGTGGCGGACCGGTATCAGGATAGCTGTGGCTACCCGTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGCGAATGGGCTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCGC 1140  
Db 1081 TTGGCGCGAATGGGCTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCGC 1140  
QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTTCCTCTGAGTTTCATGAGTTTAAACAGACCAACG 1200  
Db 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTTCCTCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGGGGATCAATTCGCGCTCTCTCCCTCCCGCCCTAACGTTACTGGC 1260  
Db 1201 GTTTCCTCTAGGGGATCAATTCGCGCTCTCTCCCTCCCGCCCTAACGTTACTGGC 1260  
QY 1261 CGAAGCGCTGGAATAAGCGGCTGTGCTTATGATGATTTTCCACCATTTG 1320  
Db 1261 CGAAGCGCTTGGATAAGCGGCTGTGCTTATGATGATTTTCCACCATTTG 1320  
QY 1321 CGGTCTTTTGGCAATGTAGGGCGCGGAAACCTGGCGCTCTCTTGTGACGAGCATTCCT 1380  
Db 1321 CGGTCTTTTGGCAATGTAGGGCGCGGAAACCTGGCGCTCTCTTGTGACGAGCATTCCT 1380  
QY 1381 AGGGTCTTTTCCCTCTCGCCAAAGAAATGCAAGGTCTGTTGAATGTCGTGAAGAGCA 1440  
Db 1381 AGGGTCTTTTCCCTCTCGCCAAAGAAATGCAAGGTCTGTTGAATGTCGTGAAGAGCA 1440  
QY 1441 GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTGAGCACCTTTGAGGAGCGG 1500  
Db 1441 GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTGAGCACCTTTGAGGAGCGG 1500

QY 1501 AACCCCCCCTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGTTGTAAGATACACT 1560  
Db 1501 AACCCCCCCTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGTTGTAAGATACACT 1560  
QY 1561 GCAAGGGGCGACAAACCCAGTGCACGTTGTGAGTTGAGTAGTTGTGGAAGAGTCAAA 1620  
Db 1561 GCAAGGGGCGACAAACCCAGTGCACGTTGTGAGTTGAGTAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAGGTACCCCATGT 1680  
Db 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAGGTACCCCATGT 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTAGTCGAGTTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTAGTCGAGTTAAAA 1740  
QY 1741 AACGCTAGGCGCCCGGAAACACGCGGAGCGTGTTCCTTTGAAAAACAGATAATACC 1800  
Db 1741 AACGCTAGGCGCCCGGAAACACGCGGAGCGTGTTCCTTTGAAAAACAGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGCGCTATTACGGCTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCTCTACAGGCGGAGACAGGAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTCTACAGGCGGAGACAGGAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGGGAGCTGGTCAATGGGCTGTGTTGGACTGTCTATCATGTGGCC 1980  
Db 1921 ACACATCTTTCTGGGAGCTGGTCAATGGGCTGTGTTGGACTGTCTATCATGTGGCC 1980  
QY 1981 GGCTCAAGACCCCTTCCCGGCCCAAGGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCCCTTCCCGGCCCAAGGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACTCTGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCATGCACTGC 2100  
Db 2041 CAGGACTCTGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCATGCACTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGTTCAGAGGATGCGGATGTCAATTCGGTGGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGTTCAGAGGATGCGGATGTCAATTCGGTGGCGCGCGG 2160  
QY 2161 GCGCAGCAGGAGGAGCTACTCTCCCGCAGCGCGCTCTCTACTTTGAAGGCTCTTTCG 2220  
Db 2161 GCGCAGCAGGAGGAGCTACTCTCCCGCAGCGCGCTCTCTACTTTGAAGGCTCTTTCG 2220  
QY 2221 GGGGCTCCTGCTCTGCGCTCGGGGACGCTGTGGGCACTTTTCGGGCTGCGGCTGTC 2280  
Db 2221 GGGGCTCCTGCTCTGCGCTCGGGGACGCTGTGGGCACTTTTCGGGCTGCGGCTGTC 2280  
QY 2281 ACCCGAGGGGTTCCGAGCGGCTGACTTTGTACCGCTGAGTCTATGAAACCACTATG 2340  
Db 2281 ACCCGAGGGGTTCCGAGCGGCTGACTTTGTACCGCTGAGTCTATGAAACCACTATG 2340  
QY 2341 CGGTCCCGGCTCTTACAGGACAACTCGTCCCTCGCGCGTACCCGAGACATTCAGGTTG 2400  
Db 2341 CGGTCCCGGCTCTTACAGGACAACTCGTCCCTCGCGCGTACCCGAGACATTCAGGTTG 2400  
QY 2401 GCCCATCTACGCGCCCTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
Db 2401 GCCCATCTACGCGCCCTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
QY 2461 GCCCAGGGTATAGGTGCTTGTCTGAACCGCTCGGCGCGCAACCTAGTTTCGGG 2520  
Db 2461 GCCCAGGGTATAGGTGCTTGTCTGAACCGCTCGGCGCGCAACCTAGTTTCGGG 2520  
QY 2521 GCGTATATGTCTAAGGACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
Db 2521 GCGTATATGTCTAAGGACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
QY 2581 ACCAGGGTGGCCCGCATCACGTAATCCACCTATGCAAGTTCCTTTCGCGGCGGTGGTTCG 2640

2581.	ACACGGGTGCCCCCAATCAGCTACTCCACTATGGCAAGTTTCTTGGCGAGGTGGTTGC	2640	3661	ATGCGATCATGTTCGGCTGACCTGGAGGTGCTCAAGAGCACTGGGTGCTGTAGCGGA	3720
2641	TCTGGGGCGCCTATGACATCATATAATATGATGAGTGCCACTCAACTGACTCGACACT	2700	3721	GTCTAGCAGCTCTGGCGCGCTATTGCTGCTGACAAACAGGACAGCGTGGTCAATTTGTGGCGAGG	3780
2641	TCTGGGGCGCCTATGACATCATATAATATGATGAGTGCCACTCAACTGACTCGACACT	2700	3721	GTCTAGCAGCTCTGGCGCGCTATTGCTGCTGACAAACAGGACAGCGTGGTCAATTTGTGGCGAGG	3780
2701	ATCTGGGCAATGCGGCAAGCTCTGGACCAAGCGGAGACGGCTGGAGCGGCACTCTGTCGTG	2760	3781	ATCATCTTTGTCGGGAAAGCGGCCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTTC	3840
2701	ATCTGGGCAATGCGGCAAGCTCTGGACCAAGCGGAGACGGCTGGAGCGGCACTCTGTCGTG	2760	3781	ATCATCTTTGTCGGGAAAGCGGCCCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTTC	3840
2761	CTGCCACCGCTACGGCTCTGGGATCGGTCAACGTGCGCAATCCAAACATCTGAGGAGGTG	2820	3841	GATGAGATGGAAGATGCGCCTCACCTCCCTTACATCGACAGGGAATGCACTCGCTCGCC	3900
2761	CTGCCACCGCTACGGCTCTGGGATCGGTCAACGTGCGCAATCCAAACATCTGAGGAGGTG	2820	3841	GATGAGATGGAAGATGCGCCTCACCTCCCTTACATCGACAGGGAATGCACTCGCTCGCC	3900
2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCTGAGACCAATC	2880	3901	GAACAAATTCAGAGGCAATCGGCTTCTGCAAAACAGCCACCAAGCAAGCGGAGGCT	3960
2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCTGAGACCAATC	2880	3901	GAACAAATTCAGAGGCAATCGGCTTCTGCAAAACAGCCACCAAGCAAGCGGAGGCT	3960
2881	AAGGGGGAGGCACTTCTGCAATTTCTGCAATTCGAAGAAATGATGATGAGCTCGCCGCG	2940	3961	GCTGTCTCCCGTGGGAATCCAAAGTGGCGGACCCCTGAAAGCTTTCTGGGCGAAGCATATG	4020
2881	AAGGGGGAGGCACTTCTGCAATTTCTGCAATTCGAAGAAATGATGATGAGCTCGCCGCG	2940	3961	GCTGTCTCCCGTGGGAATCCAAAGTGGCGGACCCCTGAAAGCTTTCTGGGCGAAGCATATG	4020
2941	AAGCTGTCCGGCTCGGACTCAATGCTGTAGCAATTTACCGGGCCCTTGATATCCGCT	3000	4021	TGAAATTTTATCAGCGGATCAATAATTTTAGCAGGCTTGTCCACTCTGCTGGAAACCCC	4080
2941	AAGCTGTCCGGCTCGGACTCAATGCTGTAGCAATTTACCGGGCCCTTGATATCCGCT	3000	4021	TGAAATTTTATCAGCGGATCAATAATTTTAGCAGGCTTGTCCACTCTGCTGGAAACCCC	4080
3001	ATACCAACTAGCGGAGAGCTCAATGCTGTAGCAACGAGCGCTTAATGACGGGCTTTACC	3060	4081	GCATAGCATCATGATGCTTACAGGCTCTATCACCAGCGCGCTCACCACCCCAACAT	4140
3001	ATACCAACTAGCGGAGAGCTCAATGCTGTAGCAACGAGCGCTTAATGACGGGCTTTACC	3060	4081	GCATAGCATCATGATGCTTACAGGCTCTATCACCAGCGCGCTCACCACCCCAACAT	4140
3061	GGCGATTTGCACTGAGTATCGACTGCAATATGATGTCACCCAGACAGTGCATTCAGC	3120	4141	ACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCGCCCACTTCTCTCTCCAGCGCT	4200
3061	GGCGATTTGCACTGAGTATCGACTGCAATATGATGTCACCCAGACAGTGCATTCAGC	3120	4141	ACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCGCCCACTTCTCTCTCCAGCGCT	4200
3121	CTGGACCGGACTTCAACATTTAGACGACGACCGGTGCCAAGACGCGGTCTACCGCTCG	3180	4201	GCTTCTGCTTTGAGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGAGCTTTGGG	4260
3121	CTGGACCGGACTTCAACATTTAGACGACGACCGGTGCCAAGACGCGGTCTACCGCTCG	3180	4201	GCTTCTGCTTTGAGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGAGCTTTGGG	4260
3181	CAGCGGCGAGGAGGACTGCTAGGGGAGGATGGGCAATTTACAGTTTGTCACTCCAGGA	3240	4261	AAGGTGCTTGTGGATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGGCC	4320
3181	CAGCGGCGAGGAGGACTGCTAGGGGAGGATGGGCAATTTACAGTTTGTCACTCCAGGA	3240	4261	AAGGTGCTTGTGGATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGGCC	4320
3241	GAAAGCGGCTCGGGATGTTGATTTCTCGGTTCTGCGAGTGCTATGACGCGGGCTGT	3300	4321	TTTAAAGGTCAATGAGCGCGAGATGCTCCACCGAGGACCTGTTTAACTTCTCTCTGCT	4380
3241	GAAAGCGGCTCGGGATGTTGATTTCTCGGTTCTGCGAGTGCTATGACGCGGGCTGT	3300	4321	TTTAAAGGTCAATGAGCGCGAGATGCTCCACCGAGGACCTGTTTAACTTCTCTCTGCT	4380
3301	GCTTGTGAGCTCACGCGCGCGAGACCTCAGTTAGTTGCGGGCTTACTTAAACACA	3360	4381	ATCCTCTCCCTCTGGCGCCCTAGTCTGCTGGGGTCTGTGGCAGCATCTGCGTGGCAC	4440
3301	GCTTGTGAGCTCACGCGCGCGAGACCTCAGTTAGTTGCGGGCTTACTTAAACACA	3360	4381	ATCCTCTCCCTCTGGCGCCCTAGTCTGCTGGGGTCTGTGGCAGCATCTGCGTGGCAC	4440
3361	CCAGGCTTCCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC	3420	4441	GTGGGCGGAGGGGCTGTGAGTGAATGAAACCGGCTGATAGAGTTTCGTTTCGCTCGCG	4500
3361	CCAGGCTTCCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC	3420	4441	GTGGGCGGAGGGGCTGTGAGTGAATGAAACCGGCTGATAGAGTTTCGTTTCGCTCGCG	4500
3421	ACCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC	3480	4501	GCTTAAACCACTCTCCCCCAACGCACTATGCTGCTGAGAGGAGCGCTGAGCAGCTGTCACT	4560
3421	ACCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC	3480	4501	GCTTAAACCACTCTCCCCCAACGCACTATGCTGCTGAGAGGAGCGCTGAGCAGCTGTCACT	4560
3481	CTGGTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTCCACTCTGCTGGGAG	3540	4561	CAGATCTCTCTAGTCTTACCATCTCAGTCTGCTGAGAGGCTTACACAGTGGATCAAC	4620
3481	CTGGTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTCCACTCTGCTGGGAG	3540	4561	CAGATCTCTCTAGTCTTACCATCTCAGTCTGCTGAGAGGCTTACACAGTGGATCAAC	4620
3541	CAAAATGTGGAAGTGTCTCATACGGCTTAAAGCCCTACGCTGCAAGCGGCGCAAGCCCTGCTG	3600	4621	GAGGACTGCTCCAGCCCATGCTCCGGCTCTGGCTTAAGAGATGTTTGGGATTTGGATATGC	4680
3541	CAAAATGTGGAAGTGTCTCATACGGCTTAAAGCCCTACGCTGCAAGCGGCGCAAGCCCTGCTG	3600	4621	GAGGACTGCTCCAGCCCATGCTCCGGCTCTGGCTTAAGAGATGTTTGGGATTTGGATATGC	4680
3601	TATAGCTGGGAGCGGTTCAAAACGAGGTTACTTACACACACCCCAATACCAATACATC	3660	4681	GAGGACTGCTCCAGCCCATGCTCCGGCTCTGGCTTAAGAGATGTTTGGGATTTGGATATGC	4740
3601	TATAGCTGGGAGCGGTTCAAAACGAGGTTACTTACACACACCCCAATACCAATACATC	3660	4681	ACGGTGTGATGATTTTCAAGACCTTGGCTCGAGTCCCAAGCTCTCTGCGCGGATTCGCGGGA	4740
3661	ATGGCATGATGTCGCTACCTGAGGTCGTCAAGGACCTTGGGTGCTGGTAGCGGA	3720	4741	GTCCCTTCTTCTCATGTCAACGTTGGGTACAGGAGGCTCTCGCGGGGCGAGCGGATCATG	4800
3661	AT				



QY	4801	CAAA	CCACTG	CTGCC	CTATG	TGGAG	CA	CAGAT	CAC	CGG	ACA	TG	TGAAAA	CGG	TTCC	AT	CAGG	4861		
DB	4801	CAAA	CCACTG	CTGCC	CTATG	TGGAG	CA	CAGAT	CAC	CGG	ACA	TG	TGAAAA	CGG	TTCC	AT	CAGG	4860		
QY	4861	AT	CGTGGG	CCCTAG	AC	CTGTAG	TAA	CAC	TG	TGG	AT	TAA	CA	TTTCCC	AT	TAA	CGCGTAC	4920		
DB	4861	AT	CGTGGG	CCCTAG	AC	CTGTAG	TAA	CAC	TG	TGG	AT	TAA	CA	TTTCCC	AT	TAA	CGCGTAC	4920		
QY	4921	ACC	ACGG	CCCTG	CA	CGCC	CT	CCCC	CG	CGC	CA	AA	TAT	TCT	AGG	CG	CTGTG	CGCGGTG	4980	
DB	4921	ACC	ACGG	CCCTG	CA	CGCC	CT	CCCC	CG	CGC	CA	AA	TAT	TCT	AGG	CG	CTGTG	CGCGGTG	4980	
QY	4981	G	TCTAGG	AG	TAC	TG	AG	GT	TAC	CG	GG	TGG	G	GA	TTTCC	A	TAC	TAC	CGG	5040
DB	4981	G	TCTAGG	AG	TAC	TG	AG	GT	TAC	CG	GG	TGG	G	GA	TTTCC	A	TAC	TAC	CGG	5040
QY	5041	ACC	ACTG	ACA	CG	TAA	GTG	CC	CG	TG	T	CA	GG	TT	CCG	CC	CG	CA	AG	5100
DB	5041	ACC	ACTG	ACA	CG	TAA	GTG	CC	CG	TG	T	CA	GG	TT	CCG	CC	CG	CA	AG	5100
QY	5101	GAT	GGG	GT	CG	GT	TG	CA	CAG	TAC	GT	TCC	AG	CG	TAA	CC	CT	TAC	GG	5160
DB	5101	GAT	GGG	GT	CG	GT	TG	CA	CAG	TAC	GT	TCC	AG	CG	TAA	CC	CT	TAC	GG	5160
QY	5161	ACA	TTCT	GTG	CG	GT	CA	AT	CA	CT	GT	TGG	GT	CA	CAG	CT	CC	CA	TG	5220
DB	5161	ACA	TTCT	GTG	CG	GT	CA	AT	CA	CT	GT	TGG	GT	CA	CAG	CT	CC	CA	TG	5220
QY	5221	CCG	AC	CT	TAG	CG	TCT	CA	CT	T	CA	TG	CT	CA	CG	A	CC	CT	CC	5280
DB	5221	CCG	AC	CT	TAG	CG	TCT	CA	CT	T	CA	TG	CT	CA	CG	A	CC	CT	CC	5280
QY	5281	G	CTA	AG	CG	TAG	CG	T	GG	CG	AG	T	CT	CC	CG	CT	T	GG	CG	5340
DB	5281	G	CTA	AG	CG	TAG	CG	T	GG	CG	AG	T	CT	CC	CG	CT	T	GG	CG	5340
QY	5341	CT	G	T	CT	CG	CG	CT	TC	CT	TG	AA	GG	CA	CA	TG	CAC	TAC	CC	5400
DB	5341	CT	G	T	CT	CG	CG	CT	TC	CT	TG	AA	GG	CA	CA	TG	CAC	TAC	CC	5400
QY	5401	CT	CA	T	CG	AG	CC	AA	CT	CT	G	GG	CG	CA	GG	AG	TGG	CG	GG	5460
DB	5401	CT	CA	T	CG	AG	CC	AA	CT	CT	G	GG	CG	CA	GG	AG	TGG	CG	GG	5460
QY	5461	TC	AG	AA	TA	AG	TAG	TAA	TTT	TG	GA	CT	TT	T	CG	AG	CG	CT	CC	5520
DB	5461	TC	AG	AA	TA	AG	TAG	TAA	TTT	TG	GA	CT	TT	T	CG	AG	CG	CT	CC	5520
QY	5521	AG	GA	AT	T	CC	GT	CC	GG	CG	AG	T	CT	G	CG	AG	GT	CC	AG	5580
DB	5521	AG	GA	AT	T	CC	GT	CC	GG	CG	AG	T	CT	G	CG	AG	GT	CC	AG	5580
QY	5581	CC	CA	T	AT	TGG	CG	CG	GA	T	TAC	AA	CC	CT	CA	CT	G	T	TAG	5640
DB	5581	CC	CA	T	AT	TGG	CG	CG	GA	T	TAC	AA	CC	CT	CA	CT	G	T	TAG	5640
QY	5641	T	AG	CT	CC	CT	CA	GT	TG	TAA	CA	CG	GT	T	CC	CA	AG	CG	CG	5700
DB	5641	T	AG	CT	CC	CT	CA	GT	TG	TAA	CA	CG	GT	T	CC	CA	AG	CG	CG	5700
QY</																				

QY	5881	TCTTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCAACGGTCTTTGG	5940
Db	5881	TCTTCATGCCCCCTTTGAGGGGAGCGGGGATCCGATCTCAGCAACGGTCTTTGG	5940
QY	5941	TCTACCCTAAGCGAGGGCTAGTGAGAACGTGCTCTGCTCTCTCATCATCATATGG	6000
Db	5941	TCTA CCGTAAGCGAGGGCTAGTGAGAACGTGCTCTGCTCTCATCATCATATGG	6000
QY	6001	ACAGGCGCCCTGATCACGCCATGGCTCGGAGGAACAACAGCTGCCATCAATGCACTG	6060
Db	6001	ACAGGCGCCCTGATCACGCCATGGCTCGGAGGAACAACAGCTGCCATCAATGCACTG	6060
QY	6061	AGCAACTCTTTGCTCCGTCAACAACCTTGTCTATGCTAAACATCTCGCAGCGCAAGC	6120
Db	6061	AGCAACTCTTTGCTCCGTCAACAACCTTGTCTATGCTAAACATCTCGCAGCGCAAGC	6120
QY	6121	CTGCGGCAGAAGAAGETCACCTTTGACAGACTCGAGGTCTGACGACACTACCGGGAC	6180
Db	6121	CTGCGGCAGAAGAAGTCACTTTGACAGACTCGAGGTCTGACGACACTACCGGGAC	6180
QY	6181	GTGCTCAAGGAGATGAAGGGCAAGCGCTCCACAGTTAAGCTAAACTTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAAGGAGATGAAGGGCAAGCGTCCACAGTTAAGGCTAAACTTTCTATCCGTGGAG	6240
QY	6241	GAAAGCTTAAGCTGACGCCCCCACATTGCGGCAGATCTAAAATTGCTATGGGGCAAAG	6300
Db	6241	GAAAGCTTAAGCTGACGCCCCCCACATTGCGGCAGATCTAAAATTGCTATGGGGCAAAG	6300
QY	6301	GACGTCGGMACTTATCAGCAAGCGCTTAAACACATCCGCTCCGTGTGGAAGGACTTGG	6360
Db	6301	GACGTCGGMACTTATCAGCAAGCGCTTAAACACATCCGCTCCGTGTGGAAGGACTTGG	6360
QY	6361	CTGGAAGACACTGAGACACCAATTGACACCAATCATGGCAAAAAATGAGGTTTTCTGC	6420
Db	6361	CTGGAAGACACTGAGACACCAATTGACACCAATCATGGCAAAAAATGAGGTTTTCTGC	6420
QY	6421	GTCCAAACAGAAAGGGGGCGCAGCCAGCTCGCTTATCGTATPCCAGATTGGGG	6480
Db	6421	GTCCAAACAGAAAGGGGGCGCAGCCAGCTCGCTTATCGTATPCCAGATTGGGG	6480
QY	6481	GTTCTGTGTGCGAGAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
Db	6481	GTTCTGTGTGCGAGAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
QY	6541	ATGGGCTCTTATACGGAITCCAATACTCTCTGGACAGCGGTGAGTTCTGTTGAAT	6600
Db	6541	ATGGGCTCTTATACGGAITCCAATACTCTCTGGACAGCGGTGAGTTCTGTTGAAT	6600
QY	6601	GCCTGNAACGGAAGAAATGCCCTATGGGCTTGCATATGACACCGCTGTTTGACTCA	6660
Db	6601	GCCTGNAACGGAAGAAATGCCCTATGGGCTTGCATATGACACCGCTGTTTGACTCA	6660
QY	6661	ACGGTCACTCAGAAATGACATCCGTTGTAGGAGTCAATCTACCAATGTTGACTTGGCC	6720
Db	6661	ACGGTCACTCAGAAATGACATCCGTTGTAGGAGTCAATCTACCAATGTTGACTTGGCC	6720
QY	6721	CCCGAAGCCAGACAGGCCATAAGTTCGCTCAAGAGCGGCTTATCATCGGGGGCCCCCTG	6780
Db	6721	CCCGAAGCCAGACAGGCCATAAGTTCGCTCAAGAGCGGCTTATCATCGGGGGCCCCCTG	6780
QY	6781	ACTAAATCTTAAGGGCAGAACTGGCGGTATCGCGGTCCCGGAGCGGTGTACTGACG	6840
Db	6781	ACTAAATCTTAAGGGCAGAACTGGCGGTATCGCGGTCCCGGAGCGGTGTACTGACG	6840
QY	6841	ACCAGCTCGGTAAATACCTCACTGTTACTTGAAGCGCGCTGCGGCTGTGAGCTTGGC	6900
Db	6841	ACCAGCTCGGTAAATACCTCACTGTTACTTGAAGCGCGCTGCGGCTGTGAGCTTGGC	6900
QY	6901	AAGCTCCAGGACTGCAAGATGCTTCGTATGCGGAGACGACCTTGTCTGTTATCTGTAAGC	6960
Db	6901	AAGCTCCAGGACTGCAAGATGCTTCGTATGCGGAGACGACCTTGTCTGTTATCTGTAAGC	6960
QY	6961	GGGGGACCAAGAGGACAGCGGAGCCTACGGGCTTACAGGAGGCTATGACTAGATAC	7020



Thu Dec 9 16:40:54 2004

us-09-576-989-6.rng

Job time : 3324 secs

Db	6961	CGGGGACCCAGAGGACGAGCGAGCTACGGGCCITTCAGGAGGCTATGACTAGATAC	7020
Qy	7021	TCTGCCCCCTCTGGGGACCGCCCAACACAGATACGACTTGGAGTTGATTAACATCATGC	7080
Db	7021	TCTGCCCCCTCTGGGGACCGCCCAACACAGATACGACTTGGAGTTGATTAACATCATGC	7080
Qy	7081	TCCTCCAAATGTCTAGTCGGGACGAGTATGCACTTGGCAAAAGGTTGATCTCTACCCGT	7140
Db	7081	TCCTCCAAATGTCTAGTCGGGACGAGTATGCACTTGGCAAAAGGTTGATCTCTACCCGT	7140
Qy	7141	GACCCACACACCCCTCTGGGGCTGCTGGGAGACAGCTTAGACACACTCCAGTCAAT	7200
Db	7141	GACCCACACACCCCTCTGGGGCTGCTGGGAGACAGCTTAGACACACTCCAGTCAAT	7200
Qy	7201	TCCTGGCTAGGCAACATCATGATGATGCGGCCACTTGTGGGCAAGGATGATCTCTGATG	7260
Db	7201	TCCTGGCTAGGCAACATCATGATGATGCGGCCACTTGTGGGCAAGGATGATCTCTGATG	7260
Qy	7261	ACTCATTTCTCTCTCATCTCTCTAGCTCAGGAACAACTTGAAGCCCTAGATTGTGAG	7320
Db	7261	ACTCATTTCTCTCTCATCTCTCTAGCTCAGGAACAACTTGAAGCCCTAGATTGTGAG	7320
Qy	7321	ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTC	7380
Db	7321	ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTC	7380
Qy	7381	CATGGCCTTAGGCAATTTCACTCCANAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
Db	7381	CATGGCCTTAGGCAATTTCACTCCANAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
Qy	7441	TCATGCTCAGAAACTTGGGCTACCGCCCTTGCAGTCTGGAGATCGGGCCAGAGT	7500
Db	7441	TCATGCTCAGAAACTTGGGCTACCGCCCTTGCAGTCTGGAGATCGGGCCAGAGT	7500
Qy	7501	GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTCTTC	7560
Db	7501	GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTCTTC	7560
Qy	7561	AACTGGGCAAGTAAAGCAAGCTCAAACTCACTCCAAATCCCGGCTCGTCCCAAGTTGGAT	7620
Db	7561	AACTGGGCAAGTAAAGCAAGCTCAAACTCACTCCAAATCCCGGCTCGTCCCAAGTTGGAT	7620
Qy	7621	TTATCCAGCTGTTCTCTGCTGCTTACAGGGGGAGAGATATATACAGCCCTGCTCTCGT	7680
Db	7621	TTATCCAGCTGTTCTCTGCTGCTTACAGGGGGAGAGATATATACAGCCCTGCTCTCGT	7680
Qy	7681	GCCGACCCCGCTGGTTCTATGCTGCTTACTCTCTACTTCTCTAGGGGTAGGCATCTAT	7740
Db	7681	GCCGACCCCGCTGGTTCTATGCTGCTTACTCTCTACTTCTCTAGGGGTAGGCATCTAT	7740
Qy	7741	CTACTCCCAACCGATGACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT	7800
Db	7741	CTACTCCCAACCGATGACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT	7800
Qy	7801	TTTTCCCTT	7860
Db	7801	TTTTCCCTT	7860
Qy	7861	TTTTTCTCTT	7920
Db	7861	TTTTTCTCTT	7920
Qy	7921	TAGCTGAAAGGTCCTGAGCCGCTTGTACTGACGAGAGTGTCTGATCTGGCCCTCTCTGC	7980
Db	7921	TAGCTGAAAGGTCCTGAGCCGCTTGTACTGACGAGAGTGTCTGATCTGGCCCTCTCTGC	7980
Qy	7981	AGATCAAGT 7989	
Db	7981	AGATCAAGT 7989	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:21:40 ; Search time 564 Seconds  
(without alignments)  
10068.249 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989  
Sequence: 1 gccagcccccagttgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7987.4	100.0	7989	4	US-09-539-601-10
2	7965.4	99.7	8001	4	US-09-539-601-7
3	7947.8	99.5	8001	4	US-09-539-601-22
4	7946.2	99.5	8001	4	US-09-539-601-16
5	7939.8	99.4	8001	4	US-09-539-601-28
6	7325.4	91.7	8637	4	US-09-539-601-4
7	7307.4	91.5	8649	4	US-09-539-601-13
8	7272.2	91.0	8639	4	US-10-029-907-1
9	7271.2	91.0	8638	4	US-10-029-907-24
10	7248.2	90.7	8638	4	US-10-029-907-7
11	7247.2	90.7	8638	4	US-10-029-907-25
12	7245.2	90.7	8642	4	US-10-029-907-2
13	7244	90.7	8638	4	US-10-029-907-6
14	7228.2	90.5	8648	4	US-10-029-907-5
15	7221	90.4	8643	4	US-10-029-907-4
16	6185.2	77.4	11076	4	US-09-539-601-1
17	6167.6	77.2	11076	4	US-09-539-601-25
18	6166	77.2	11076	4	US-09-539-601-19
19	6159.6	77.1	11076	4	US-09-539-601-31
20	5316.8	66.6	9595	3	US-09-014-416-4
21	5216.4	65.3	7917	1	US-08-324-977-31
22	5216.4	65.3	7917	2	US-08-384-616-31
23	5216.4	65.3	7917	2	US-08-904-686A-31
24	5216.4	65.3	7917	3	US-09-315-850-31
25	5216.4	65.3	9416	1	US-08-324-977-1
26	5216.4	65.3	9416	2	US-08-384-616-1
27	5216.4	65.3	9416	2	US-08-904-686A-1

28 5216.4 65.3 9416 3 US-09-315-850-1  
29 5216.4 65.3 9416 4 US-08-823-895A-27  
30 5210.2 65.2 9472 4 US-08-150-204B-96  
31 5175.2 64.8 7863 1 US-08-324-977-35  
32 5175.2 64.8 7863 2 US-08-384-616-35  
33 5175.2 64.8 7863 2 US-08-904-686A-35  
34 5175.2 64.8 7863 3 US-09-315-850-35  
35 5175.2 64.8 9030 1 US-08-324-977-13  
36 5175.2 64.8 9030 2 US-08-384-616-13  
37 5175.2 64.8 9030 3 US-08-904-686A-13  
38 5175.2 64.8 9030 3 US-09-315-850-13  
39 4102.2 51.3 9599 3 US-09-014-416-2  
40 4100.6 51.3 9599 3 US-09-014-416-6  
41 4096.2 51.3 9646 3 US-08-811-566-1  
42 4096.2 51.3 9646 3 US-09-034-756-1  
43 4092.6 51.2 12980 3 US-08-811-566-5  
44 4092.6 51.2 12980 3 US-09-034-756-5  
45 4019.8 50.3 9379 3 US-09-388-874-1

#### ALIGNMENTS

RESULT 1  
US-09-539-601-10  
; Sequence 10, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 7989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)...(341)  
; OTHER INFORMATION: construct 1377/NS3-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1181)  
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
; OTHER INFORMATION: fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1190)..(1800)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1801)..(7758)  
; OTHER INFORMATION: hepatitis C virus NS3 - 5B  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (7759)..(7989)  
; PUBLICATION INFORMATION:  
; AUTHORS: Lohmann, Volker  
; AUTHORS: Kner, Frank  
; AUTHORS: Koch, Jan-Oliver  
; AUTHORS: Herian, Ulrike  
; AUTHORS: Theilmann, Lorenz  
; AUTHORS: Bartenschlager, Ralf  
; TITLE: Replication of subgenomic hepatitis C virus RNAs in a  
; TITLE: Hepatoma cell line  
; JOURNAL: Science  
; VOLUME: 285

Sequence 1, Appl  
Sequence 27, Appl  
Sequence 96, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 1, Appl

;  
; PAGES: 110-113  
; DATE: 1999-07-02  
; US-09-539-601-10

Query Match  
Best Local Similarity 100.0%; Score 7987.4; DB 4; Length 7989;  
Matches 7989; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATGGGGGACACTCCACATAGATCACTCCCTGTGAGGAACACTAG 60  
Db 1 GCCAGCCCCGATGGGGGACACTCCACATAGATCACTCCCTGTGAGGAACACTAG 60

QY 61 TCTTCACGCAAGAGCGTCTAGCCATGGCGTGTAGTATGAGTGTGTCAGCCTCCAGAC 120  
Db 61 TCTTCACGCAAGAGCGTCTAGCCATGGCGTGTAGTATGAGTGTGTCAGCCTCCAGAC 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAG 180  
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAG 180

QY 181 GAGCAGCGGTCTCTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGCGGTGCCCC 240  
Db 181 GAGCAGCGGTCTCTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGCGGTGCCCC 240

QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGCGCTTGTGTAATTCGCTGATAGG 300  
Db 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGCGCTTGTGTAATTCGCTGATAGG 300

QY 301 GTGCTTGGAGTGCCTCCCGGAGTCTGTAAGCCGTGACCGTACCATGAGCAGAAATCCTAAC 360  
Db 301 GTGCTTGGAGTGCCTCCCGGAGTCTGTAAGCCGTGACCGTACCATGAGCAGAAATCCTAAC 360

QY 361 CTCAAGAAACCAAGGCGCGCATGATTCGCTATGCTGGGACACAGACAAATCGGTGCT 420  
Db 361 CTCAAGAAACCAAGGCGCGCATGATTCGCTATGCTGGGACACAGACAAATCGGTGCT 420

QY 421 CGSCCGCTTGGTGGAGAGCTATTCGGCTATGCTGGGACACAGACAAATCGGTGCT 480  
Db 421 CGSCCGCTTGGTGGAGAGCTATTCGGCTATGCTGGGACACAGACAAATCGGTGCT 480

QY 481 CTGATGCGCGGTTCGGCTGTCAGCGAGGGCGCGCGTCTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGGTTCGGCTGTCAGCGAGGGCGCGCGTCTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCGGTGCCTGATGAATGACTGACAGCAGCGCGCGCTATCGTGGCTGCCA 600  
Db 541 ACCTGTCGGTGCCTGATGAATGACTGACAGCAGCGCGCGCTATCGTGGCTGCCA 600

QY 601 CGACGGCGTTCCTTCGCGAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGC 660  
Db 601 CGACGGCGTTCCTTCGCGAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGC 660

QY 661 TGTATTTGGCGAGTTCGGCGGAGAGTCTCTGTCACTCTACCTTGTCTTCGCGAGA 720  
Db 661 TGTATTTGGCGAGTTCGGCGGAGAGTCTCTGTCACTCTACCTTGTCTTCGCGAGA 720

QY 721 AAGTATCCATATGCTGATGCAATCGCGGCTGCATACGCTTGTATCGGCTTACCTGCC 780  
Db 721 AAGTATCCATATGCTGATGCAATCGCGGCTGCATACGCTTGTATCGGCTTACCTGCC 780

QY 781 CATTCGACCAACGGAACATCGCATCGAGCAGCAGTACTCGGATGGAAGCGGTC 840  
Db 781 CATTCGACCAACGGAACATCGCATCGAGCAGCAGTACTCGGATGGAAGCGGTC 840

QY 841 TTGTGATCAGATGATCTGGAGAGAGATCAGGGGTTCGGCGAGCGGAATGTTTCG 900  
Db 841 TTGTGATCAGATGATCTGGAGAGAGATCAGGGGTTCGGCGAGCGGAATGTTTCG 900

QY 901 CCAGGCTCAAGCGCGCATGCCGAGAGATCTCGTGTGACCCATGCGATGCGCT 960  
Db 901 CCAGGCTCAAGCGCGCATGCCGAGAGATCTCGTGTGACCCATGCGATGCGCT 960

QY 961 GCTTCCCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGC 1020

Db 961 GCTTCCCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGC 1020

QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080

Db 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080

QY 1081 TTGGCGGCAATGGGTGACCGTTCCTGTGCTTTACGGTATCGCGCTCCCGATTGCG 1140

Db 1081 TTGGCGGCAATGGGTGACCGTTCCTGTGCTTTACGGTATCGCGCTCCCGATTGCG 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGTTAAACAGACCAACAGC 1200

Db 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGTTAAACAGACCAACAGC 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCTCCCGCTAACTGTTACTGGC 1260

Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCTCCCGCTAACTGTTACTGGC 1260

QY 1261 CGAAGCGCTTGAATAAGGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG 1320

Db 1261 CGAAGCGCTTGAATAAGGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAACCTGGCTCTTCTTGAAGAGCATTCCT 1380

Db 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAACCTGGCTCTTCTTGAAGAGCATTCCT 1380

QY 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGTCTGTGTAATGTGTGAAGGAGCA 1440

Db 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGTCTGTGTAATGTGTGAAGGAGCA 1440

QY 1441 GTTTCCTCTGGAAGTCTTGAAGACAAACAACTGTAGCGACCTTTTCAGGCGCGG 1500

Db 1441 GTTTCCTCTGGAAGTCTTGAAGACAAACAACTGTAGCGACCTTTTCAGGCGCGG 1500

QY 1501 AACCCCGCTCTGGGACAGGTCTCTCGGCAAAAGCAAGCTGTGTAAGGATACACCT 1560

Db 1501 AACCCCGCTCTGGGACAGGTCTCTCGGCAAAAGCAAGCTGTGTAAGGATACACCT 1560

QY 1561 GCAAGGCGGCAACACCCAGTSCCAGTGTGAGTGTGATAGTTGGAAGAGTCAAA 1620

Db 1561 GCAAGGCGGCAACACCCAGTSCCAGTGTGAGTGTGATAGTTGGAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCTATTCACAGGGGCTGAAGGATGCCAGAGTACCCATTTG 1680

Db 1621 TGGCTCTCTCAAGCTATTCACAGGGGCTGAAGGATGCCAGAGTACCCATTTG 1680

QY 1681 ATGGGATCTGATCTGGGCTCGGTGCATGCTTTACATGCTTTAGTGTGAGTTTAAA 1740

Db 1681 ATGGGATCTGATCTGGGCTCGGTGCATGCTTTACATGCTTTAGTGTGAGTTTAAA 1740

QY 1741 AACGTCAGGCGGCAACACCCAGTSCCAGTGTGAGTGTGATAGTTGGAAGAGTCAAA 1800

Db 1741 AACGTCAGGCGGCAACACCCAGTSCCAGTGTGAGTGTGATAGTTGGAAGAGTCAAA 1800

QY 1801 ATGGGCTCTTATTCGCTCTCTCCAGAGCGGCTTACTTGGCTGCATCATCT 1860

Db 1801 ATGGGCTCTTATTCGCTCTCTCCAGAGCGGCTTACTTGGCTGCATCATCT 1860

QY 1861 AGCCTCACAGGCGGCAACAGGTCAGGGGAGGTCCAGTGTGCTCCACCGCA 1920

Db 1861 AGCCTCACAGGCGGCAACAGGTCAGGGGAGGTCCAGTGTGCTCCACCGCA 1920

QY 1921 ACACATCTTCTCGCGACCTGCGTCAATGGCGTGTGTGAGTGTCTATCATGTGTC 1980

Db 1921 ACACATCTTCTCGCGACCTGCGTCAATGGCGTGTGTGAGTGTCTATCATGTGTC 1980

QY 1981 GGTCAAGACCTTTCGCGGCGCAACAGGCGGCTTACTTGGCTGCATCATCT 2040

Db 1981 GGTCAAGACCTTTCGCGGCGCAACAGGCGGCTTACTTGGCTGCATCATCT 2040

QY 2041 CAGGACCTCTGTCGGTGGCAAGCGCGGCGGTTCCTTGAACCATGCACTGCTGC 2100

Db 2041 CAGGACCTCGTCGGCTGGCAAGCGCCCGCGCGGCTTCTCTTGACACCATGACACCTGC 2100  
Qy 2101 GGCAGCTCGGACCTTACTTGGTACAGAGGATGCGGATGTCATTCGGTGGCGCGGG 2160  
Db 2101 GGCAGCTCGGACCTTACTTGGTACAGAGGATGCGGATGTCATTCGGTGGCGCGGG 2160  
Qy 2161 GGCACAGCAGGGGAGGCTACTCTCCCGCAGGCGGCTCTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGCACAGCAGGGGAGGCTACTCTCCCGCAGGCGGCTCTCTCTACTTGAAGGGCTCTTCG 2220  
Qy 2221 GCGGTCCACTGCTCTGCTCCCTCGGGGACGCTGTGGGCATCTTTTCGGGCTGCGGTGTC 2280  
Db 2221 GCGGTCCACTGCTCTGCTCCCTCGGGGACGCTGTGGGCATCTTTTCGGGCTGCGGTGTC 2280  
Qy 2281 ACCCGAGGGGTGGAGGGCGGTGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCCGAGGGGTGGAGAGGGCGGTGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Qy 2341 CCGTCCCGGCTTTCAGGGACAACTCGTCCCTCGGGCGGTACCGCAGACATTCAGGTG 2400  
Db 2341 CCGTCCCGGCTTTCAGGGACAACTCGTCCCTCGGGCGGTACCGCAGACATTCAGGTG 2400  
Qy 2401 GCCCATCTACAGCCCTACTGTGTAGGGCAAGAGCACTAAGGTGCCGGCTGGGTATGCA 2460  
Db 2401 GCCCATCTACAGCCCTACTGTGTAGGGCAAGAGCACTAAGGTGCCGGCTGGGTATGCA 2460  
Qy 2461 GCCCAAGGGTATAGGTGCTTGTCTGAAACCGTCCGTGCGCGGCACTAGGTTCGGG 2520  
Db 2461 GCCCAAGGGTATAGGTGCTTGTCTGAAACCGTCCGTGCGCGGCACTAGGTTCGGG 2520  
Qy 2521 GCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
Db 2521 GCGTATATGCTTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
Qy 2581 ACCACGGGTGCCCGCATCACTACTCCACCTATGCGCAAGTTCCTTCCGACGGTGTTC 2640  
Db 2581 ACCACGGGTGCCCGCATCACTACTCCACCTATGCGCAAGTTCCTTCCGACGGTGTTC 2640  
Qy 2641 TCTGGGGCGCCTATGACATCATATATGATGAGTGCGACTCACTCACTGACTGACCACT 2700  
Db 2641 TCTGGGGCGCCTATGACATCATATATGATGAGTGCGACTCACTCACTGACTGACCACT 2700  
Qy 2701 ATCTGGGCATCGGCACAGTCTCGACCAAGCGAGACGGTGGAGCGGACTCGTGTG 2760  
Db 2701 ATCTGGGCATCGGCACAGTCTCGACCAAGCGAGACGGTGGAGCGGACTCGTGTG 2760  
Qy 2761 CTGCCACCGCTAGCCTCGGGATCGGTCAACGTCGCACATCCAAACATCGAGAGGTG 2820  
Db 2761 CTGCCACCGCTAGCCTCGGGATCGGTCAACGTCGCACATCCAAACATCGAGAGGTG 2820  
Qy 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCGATCCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCGATCCCCATCGAGACCATC 2880  
Qy 2881 AAGGGGGAGGACCTCACTTTCTGCCATTCGAAGAAAGTGTGATGAGTCCGCGG 2940  
Db 2881 AAGGGGGAGGACCTCACTTTCTGCCATTCGAAGAAAGTGTGATGAGTCCGCGG 2940  
Qy 2941 AAGCTGTCCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATGATCGTC 3000  
Db 2941 AAGCTGTCCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATGATCGTC 3000  
Qy 3001 ATACCACTAGCGAGAGCTCATGTGTAGCAACGAGCGCTCATATGAGCGGCTTTACC 3060  
Db 3001 ATACCACTAGCGAGAGCTCATGTGTAGCAACGAGCGCTCATATGAGCGGCTTTACC 3060  
Qy 3061 GCGGATTTGACTCAGTGATCGACTCAATACATGTGTCCAGCAGAGTCTGATTCAGC 3120  
Db 3061 GCGGATTTGACTCAGTGATCGACTCAATACATGTGTCCAGCAGAGTCTGATTCAGC 3120  
Qy 3121 CTGACCCGACCTTACCAATTGAGACGACGACCGGTGCGCAAGACGGGTGTCAACGCTG 3180  
Db 3121 CTGACCCGACCTTACCAATTGAGACGACGACCGGTGCGCAAGACGGGTGTCAACGCTG 3180

Qy 3181 CAGCGGCGAGGACAGACTGTGTAGGGGAGAGATGGGCATTTTACAGTTTGTGACTCCAGG 3240  
Db 3181 CAGCGGCGAGGACAGACTGTGTAGGGGAGAGATGGGCATTTTACAGTTTGTGACTCCAGG 3240  
Qy 3241 GAAAGGGCTCGGGCATGTTCCGATTCTCGTTCTGTGCGAGTCTATGACGCGGCTGT 3300  
Db 3241 GAAAGGGCTCGGGCATGTTCCGATTCTCGTTCTGTGCGAGTCTATGACGCGGCTGT 3300  
Qy 3301 GCTTGTGTACAGAGTTCAGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACTTAAACACA 3360  
Db 3301 GCTTGTGTACAGAGTTCAGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACTTAAACACA 3360  
Qy 3361 CGAGGTGGCCGCTGTCAGAGGACCATCTGAGTTCTGGAGAGCGTCTTTTACAGGCTC 3420  
Db 3361 CGAGGTGGCCGCTGTCAGAGGACCATCTGAGTTCTGGAGAGCGTCTTTTACAGGCTC 3420  
Qy 3421 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGACAACTTCCCTTAC 3480  
Db 3421 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGAGAACTTCCCTTAC 3480  
Qy 3481 CTGTGTACATACAGAGTTCAGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
Db 3481 CTGTGTACATACAGAGTTCAGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
Qy 3541 CAATGTGGAGTGTCTCATACGCTAAAGCTAAGCTGCAAGGGGCAACGCCCTGCTG 3600  
Db 3541 CAATGTGGAGTGTCTCATACGCTAAAGCTAAGCTGCAAGGGGCAACGCCCTGCTG 3600  
Qy 3601 TATAGGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCATTAACCAATATATC 3660  
Db 3601 TATAGGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCATTAACCAATATATC 3660  
Qy 3661 ATGCGATGATGCGGTGCTGACCTGGAGTCTGTCCAGAGACCTGGGTGCTGTAGGCGGA 3720  
Db 3661 ATGCGATGATGCGGTGCTGACCTGGAGTCTGTCCAGAGACCTGGGTGCTGTAGGCGGA 3720  
Qy 3721 GTCTTAGCAGCTGTGGCGGCTATGCTGTGCAACAGGAGCGGTGCTATTTGGGAGG 3780  
Db 3721 GTCTTAGCAGCTGTGGCGGCTATGCTGTGCAACAGGAGCGGTGCTATTTGGGAGG 3780  
Qy 3781 ATCATCTTGTCCGGAAGCGGCACTATTTCCGAGAGGAACTCTTTTACGGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGGAAGCGGCACTATTTCCGAGAGGAACTCTTTTACGGGAGTTC 3840  
Qy 3841 GATGAGATGGAAGTGTGCGCTCACTCTCCCTTACATCGAAAGGAAATGACGCTGCC 3900  
Db 3841 GATGAGATGGAAGTGTGCGCTCACTCTCCCTTACATCGAAAGGAAATGACGCTGCC 3900  
Qy 3901 GAAACAATTCAAAAGAGGCAATCGGGTTCGTCAAAACAGCCCAAGCAAGCGAGGCT 3960  
Db 3901 GAAACAATTCAAAAGAGGCAATCGGGTTCGTCAAAACAGCCCAAGCAAGCGAGGCT 3960  
Qy 3961 GCTGCTCCGCTGTGGAATCCAAAGTGGGACCGCTCGAGCCTTCTGGGAGAGCATATG 4020  
Db 3961 GCTGCTCCGCTGTGGAATCCAAAGTGGGACCGCTCGAGCCTTCTGGGAGAGCATATG 4020  
Qy 4021 TGGAAATTTTCATAGCGGGATACAAATTTTACAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTTTCATAGCGGGATACAAATTTTACAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Qy 4081 GCGATAGATCATGATGAGCATTCACAGCTCTATCAACAGCCCGCTCAACCCCAACAT 4140  
Db 4081 GCGATAGATCATGATGAGCATTCACAGCTCTATCAACAGCCCGCTCAACCCCAACAT 4140  
Qy 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGCCCACTTGTCTCTCCAGGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGCCCACTTGTCTCTCCAGGCT 4200  
Qy 4201 GCTTCTGCTTTCGTAGCGCGGCACTCGCTGGAGCGGCTGTTGGCAGATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGCGCGGCACTCGCTGGAGCGGCTGTTGGCAGATAGGCTTGGG 4260



Db 6421 GTCCACAGAGAGAGGGGGCCGCAAGCAGCTCGCCTTATGTTATCCAGATTTGGG 6480  
Qy 6481 GTTCGTGTGTCGAGAAATGGCCCTTTACGATGTGTCTCCACCCCTCCCTCAGGCCGTG 6540  
Db 6481 GTTCGTGTGTCGAGAAATGGCCCTTTACGATGTGTCTCCACCCCTCCCTCAGGCCGTG 6540  
Qy 6541 ATGGGCTCTCATACGGAATCCATATCTCTCCCTGGACAGGGGTCCGAGTTCCTGTTGAAT 6600  
Db 6541 ATGGGCTCTCATACGGAATCCATATCTCTCCCTGGACAGGGGTCCGAGTTCCTGTTGAAT 6600  
Qy 6601 GCCTGGAAGCGGAAGAAATGCCCTATCGGCTTCGATATGACACCCCGCTGTTTGACTCA 6660  
Db 6601 GCCTGGAAGCGGAAGAAATGCCCTATCGGCTTCGATATGACACCCCGCTGTTTGACTCA 6660  
Qy 6661 ACGTCTACTGAGAATGACATCCGCTGTTGAGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Db 6661 ACGTCTACTGAGAATGACATCCGCTGTTGAGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Qy 6721 CCGGAGCCAGACAGGCCATAAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
Db 6721 CCGGAGCCAGACAGGCCATAAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
Qy 6781 ACTAATCTTAAAGCGAGAACTGCGGCTATCGCGGTGCGCGAGCGGTGTACTGACG 6840  
Db 6781 ACTAATCTTAAAGCGAGAACTGCGGCTATCGCGGTGCGCGAGCGGTGTACTGACG 6840  
Qy 6841 ACCAGTCGCGGTAAATACCCCTCAATGTTACTTGAAGCGCGCTGGCGCTGTGAGCTGGC 6900  
Db 6841 ACCAGTCGCGGTAAATACCCCTCAATGTTACTTGAAGCGCGCTGGCGCTGTGAGCTGGC 6900  
Qy 6901 AAGCTCAGACTGCACGATGCTCGTATCGGAGACGACCTTGTCTGTGAAAGC 6960  
Db 6901 AAGCTCAGACTGCACGATGCTCGTATCGGAGACGACCTTGTCTGTGAAAGC 6960  
Qy 6961 GCGGGACCCAAAGAGGAGGAGGAGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAAAGAGGAGGAGGAGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Qy 7021 TCTGCCCTTGGGAGACCGCCCAACACAGATACGACTTGGAGTTGATCAATCATGTC 7080  
Db 7021 TCTGCCCTTGGGAGACCGCCCAACACAGATACGACTTGGAGTTGATCAATCATGTC 7080  
Qy 7081 TCCTCCAAATGTTCAGTCCGCGAGTGCATCTGGCAAAAGGTTACTATCTCACCCGT 7140  
Db 7081 TCCTCCAAATGTTCAGTCCGCGAGTGCATCTGGCAAAAGGTTACTATCTCACCCGT 7140  
Qy 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
Db 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
Qy 7201 TCCTGGCTAGGCAACATCATGTTATGCGGCCACCTTGGGCAAGGATGATCCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATGTTATGCGGCCACCTTGGGCAAGGATGATCCTGATG 7260  
Qy 7261 ACTCATTTCTTCCATCTCTAGCTCAGAACAACTTGAAGAAGCCCTAGATTGTTCAG 7320  
Db 7261 ACTCATTTCTTCCATCTCTAGCTCAGAACAACTTGAAGAAGCCCTAGATTGTTCAG 7320  
Qy 7321 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCAATCAACGACTC 7380  
Db 7321 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCAATCAACGACTC 7380  
Qy 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCAGGTGAGATCAATAGGGTGGCT 7440  
Db 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCAGGTGAGATCAATAGGGTGGCT 7440  
Qy 7441 TCATGCTCAGGAACCTTGGGTACCGCTTGGAGTCTGGACATCGCGCCAGAGT 7500  
Db 7441 TCATGCTCAGGAACCTTGGGTACCGCTTGGAGTCTGGAGACATCGCGCCAGAGT 7500  
Qy 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC 7560  
Db 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC 7560

## RESULT 2

US-09-539-601-7  
; Sequence 7, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/NS3-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core-neomycin  
; OTHER INFORMATION: phosphotransferase fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1202)..(1812)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1813)..(7770)  
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3-5B  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (7771)..(8001)

PUBLICATION INFORMATION:  
; AUTHORS: Lohmann, Volker  
; AUTHORS: Krner, Frank  
; AUTHORS: Koch, Jan-Oliver  
; AUTHORS: Herian, Ulrike  
; AUTHORS: Theilmann, Lorenz  
; AUTHORS: Bartenschlager, Ralf  
; TITLE: Replication of subgenomic hepatitis c virus RNAs in a  
; TITLE: hepatoma cell line  
; JOURNAL: Science  
; VOLUME: 285  
; PAGES: 110-113  
; DATE: 1999-07-02  
US-09-539-601-7

Query Match 99.7%; Score 7965.4; DB 4; Length 8001;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 7988; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

QY	1	GCAGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACCTACTG	60
DB	1	GCAGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACCTACTG	60
QY	61	TCTTCAACGAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGTCAGCCTCCAGGAC	120
DB	61	TCTTCAACGAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGTCAGCCTCCAGGAC	120
QY	121	CCCCCTCCGGAGAGCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTGCCAG	180
DB	121	CCCCCTCCGGAGAGCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTGCCAG	180
QY	181	GACGACCGGTCCTTCTTGGATCAACCCGCTCAATGCTCGAGATTTCGGGTCGCCCC	240
DB	181	GACGACCGGTCCTTCTTGGATCAACCCGCTCAATGCTCGAGATTTCGGGTCGCCCC	240
QY	241	GGGAGCTGTACCGGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTGCTGATAGG	300
DB	241	GGGAGCTGTACCGGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTGCTGATAGG	300
QY	301	GTGCTTGGAGTCCCGGGAGTCTCGTAGCCGTCGACCATGACGAGCAATCTTAAC	360
DB	301	GTGCTTGGAGTCCCGGGAGTCTCGTAGCCGTCGACCATGACGAGCAATCTTAAC	360
QY	361	CTCAAGAAAAACCAAA-----GGGCGGCCCATGATTGAACAAGATGGATTGC	408
DB	361	CTCAAGAAAAACCAAACTAACAACCGGGCGCCATGATTGAACAAGATGGATTGC	420
QY	409	ACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTGGCTATGACTGGGCACACAGA	468
DB	421	ACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTGGCTATGACTGGGCACACAGA	480
QY	469	CAATCGGCTGTGATGCGCGGTGTTCCGGCTGTACGAGCGGGGCGCCGGTCTTT	528
DB	481	CAATCGGCTGTGATGCGCGGTGTTCCGGCTGTACGAGCGGGGCGCCGGTCTTT	540
QY	529	TTGTCAAGACCGACCTGTCCGGTCCCTGAATGAACCTGACGACGAGGCGCGGCTAT	588
DB	541	TTGTCAAGACCGACCTGTCCGGTCCCTGAATGAACCTGACGACGAGGCGCGGCTAT	600
QY	589	CGTGCTGGCCACGACGCGGCTTCTTGGCGAGTGTGTCAGCGTGTCACTGAAGCGG	648
DB	601	CGTGCTGGCCACGACGCGGCTTCTTGGCGAGTGTGTCAGCGTGTCACTGAAGCGG	660
QY	649	GAAGGACTGGCTGATTGGGCGAGTGC CGGGCAGGATCTCTGTCATCTCACTTG	708
DB	661	GAAGGACTGGCTGATTGGGCGAGTGC CGGGCAGGATCTCTGTCATCTCACTTG	720
QY	709	CTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGGCGGCTGATCGCTGATC	768
DB	721	CTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGGCGGCTGATCGCTGATC	780
QY	769	CGGCTACCTGCCCATTCGACCAACCAAGCAAAATCGATCGAGGAGCGACTCTCGGA	828
DB		CGGCTACCTGCCCATTCGACCAACCAAGCAAAATCGATCGAGGAGCGACTCTCGGA	

DB	781	CGGCTACCTGCCCATTCGACCAACCAAGCAAAATCGATCGAGCGAGCACTCTCGGA	840
QY	829	TGGAAGCCGGTCTTGTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAG	888
DB	841	TGGAAGCCGGTCTTGTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAG	900
QY	889	CGGAAGCTGTCGCCAGGCTCAAGGCGCATCCGAGCGGAGGATCTCGTCTGACCC	948
DB	901	CGGAAGCTGTCGCCAGGCTCAAGGCGCATCCGAGCGGAGGATCTCGTCTGACCC	960
QY	949	ATGGCGATGCTGCTTGGCGAATATCATGTTGGAATATGCGCGCTTTCTGATTCATCG	1008
DB	961	ATGGCGATGCTGCTTGGCGAATATCATGTTGGAATATGCGCGCTTTCTGATTCATCG	1020
QY	1009	ACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGACATAGCTTGGTACCGGTGATA	1068
DB	1021	ACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGACATAGCTTGGTACCGGTGATA	1080
QY	1069	TTGCTGAAGAGCTTGGCGGCAATGGCTGACCGTCTCTGCTTACGGTATCGCG	1128
DB	1081	TTGCTGAAGAGCTTGGCGGCAATGGCTGACCGTCTCTGCTTACGGTATCGCG	1140
QY	1129	CTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTTTAA	1188
DB	1141	CTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTTTAA	1200
QY	1189	GAGACCAACGCTTTCCTCTAGCGGATCAATTCGCGCTCTTCCCTCCCGCCCT	1248
DB	1201	GAGACCAACGCTTTCCTCTAGCGGATCAATTCGCGCTCTTCCCTCCCGCCCT	1260
QY	1249	AACGTTACTGGCGAAGCGCTTGAATAAGCGCGTGTGCTTGTCTATATGTTAT	1308
DB	1261	AACGTTACTGGCGAAGCGCTTGAATAAGCGCGTGTGCTTGTCTATATGTTAT	1320
QY	1309	TCCACCAATATGCGCTTGTGGCAATGTAGGGCCCGGAAACCTGGCTGTCTTGT	1368
DB	1321	TCCACCAATATGCGCTTGTGGCAATGTAGGGCCCGGAAACCTGGCTGTCTTGT	1380
QY	1369	ACGAGCTTCTAGGGTCTTCCCTCTGCGCAAGGATCAAGGCTGTGTGATGTC	1428
DB	1381	ACGAGCTTCTAGGGTCTTCCCTCTGCGCAAGGATCAAGGCTGTGTGATGTC	1440
QY	1429	GTGAAGGAGCAGATTCTCTGGAAGCTTCTTGAAGACAAACCAACCTGTAGCGACCT	1488
DB	1441	GTGAAGGAGCAGATTCTCTGGAAGCTTCTTGAAGACAAACCAACCTGTAGCGACCT	1500
QY	1489	TGCGAGCGGAAACCCCGACCTGGCGACAGTGTCTGCGGCAAAAGCCACGTGTA	1548
DB	1501	TGCGAGCGGAAACCCCGACCTGGCGACAGTGTCTGCGGCAAAAGCCACGTGTA	1560
QY	1549	TAAGATACCTGCAAGCGGCAACCCCGACCTGGCGACAGTGTGTGATGTTGTG	1608
DB	1561	TAAGATACCTGCAAGCGGCAACCCCGACCTGGCGACAGTGTGTGATGTTGTG	1620
QY	1609	GAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGCTCAAGGATCCACAGAG	1668
DB	1621	GAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGCTCAAGGATCCACAGAG	1680
QY	1669	GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGACATGCTTACATGTTTAT	1728
DB	1681	GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGACATGCTTACATGTTTAT	1740
QY	1729	TCGAGTTTAAAAACGCTTAGGCCCGGCAACCGGAGCGTGTCTTCTTCAAAA	1788
DB	1741	TCGAGTTTAAAAACGCTTAGGCCCGGCAACCGGAGCGTGTCTTCTTCAAAA	1800
QY	1789	CACGATAATCATGGCGCTTATACGGCTACTCTCCCAACGACGCGAGGCTACTTGC	1848
DB	1801	CACGATAATCATGGCGCTTATACGGCTACTCTCCCAACGACGCGAGGCTACTTGC	1860
QY	1849	TGCATCATCTAGCTCAGGCGCGGACAGGAACCGAGTCTGAGGGGAGGTCCAGTG	1908
DB	1861	TGCATCATCTAGCTCAGGCGCGGACAGGAACCGAGTCTGAGGGGAGGTCCAGTG	1920



1909 GTCTCCACCGCAACAATCTTTCTCTGGCGAAGCTGGTCAATGGCGTGTGGAGCTGTC 1968  
1921 GTCTCCACCGCAACAATCTTTCTCTGGCGAAGCTGGTCAATGGCGTGTGGAGCTGTC 1980  
1969 TATCATGGTCCGGTCAAGACCTTTGCGGCCCAAGGGCCCAATACCCAAATGTAC 2028  
1981 TATCATGGTCCGGTCAAGACCTTTGCGGCCCAAGGGCCCAATACCCAAATGTAC 2040  
2029 ACCAATGTGGACAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACA 2088  
2041 ACCAATGTGGACAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACA 2100  
2089 CGATGACCTCGCGGAGCTCGGACCTTTATTTGGTCAAGAGGATCGGATGTCATCCG 2148  
2101 CGATGACCTCGCGGAGCTCGGACCTTTATTTGGTCAAGAGGATCGGATGTCATCCG 2160  
2149 GTGCGCGCGGGGACAGAGGGGGAGCTACTCTCCCAAGGCCGCTCTCTACTTTG 2208  
2161 GTGCGCGCGGGGACAGAGGGGGAGCTACTCTCCCAAGGCCGCTCTCTACTTTG 2220  
2209 AAGGGCTCTTTGGGGGCTCCACTGCTCTGCCCCCTCGGGGCAAGCTGTGGGCAATCTTCG 2268  
2221 AAGGGCTCTTTGGGGGCTCCACTGCTCTGCCCCCTCGGGGCAAGCTGTGGGCAATCTTCG 2280  
2269 GCTGCGGTGTGACACCGAGGGGTTCGAGGGCGGTGGAATTTGTACCGTCGAGTCTATG 2328  
2281 GCTGCGGTGTGACACCGAGGGGTTCGAGGGCGGTGGAATTTGTACCGTCGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGGCGGTACCGCAG 2388  
2341 GAAACCACTATGCGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGGCGGTACCGCAG 2400  
2389 ACATTCAGGTGCCCCATCTACAGCCCCCTACTGGTAGCGGCAAGAGCACTAAGGTGCCG 2448  
2401 ACATTCAGGTGCCCCATCTACAGCCCCCTACTGGTAGCGGCAAGAGCACTAAGGTGCCG 2460  
2449 GCTGCGTATGACGCCAAGGGTATAGGTGCTTGTCTGAAACCGCTCGTGGCGGCCACC 2508  
2461 GCTGCGTATGACGCCAAGGGTATAGGTGCTTGTCTGAAACCGCTCGTGGCGGCCACC 2520  
2509 CTAGGTTCTGGGGGTATATGCTTAGGGACATGGTATCGACCTAACATCAGAACCGGG 2568  
2521 CTAGGTTCTGGGGGTATATGCTTAGGGACATGGTATCGACCTAACATCAGAACCGGG 2580  
2569 GTAAGGACCATCACCGGTGCCCCCATCACGTACTTCCACCTATGGCAAGTTTCTTGCC 2628  
2581 GTAAGGACCATCACCGGTGCCCCCATCACGTACTTCCACCTATGGCAAGTTTCTTGCC 2640  
2629 GACGGTGGTGTCTGCGGGGCGCTATGACATCATATATGATGATGATGATGATGATGAT 2688  
2641 GACGGTGGTGTCTGCGGGGCGCTATGACATCATATATGATGATGATGATGATGATGAT 2700  
2689 GACTCGACCACTATCTGGGCACTGGCACAGTCTCGACCAAGCGGAGCGGTGGAGCG 2748  
2701 GACTCGACCACTATCTGGGCACTGGCACAGTCTCGACCAAGCGGAGCGGTGGAGCG 2760  
2749 CGACTCGTGTGTCGACCGGTACCGCTCCGGGATCGGTCAAGTGGCCACATCCAAAC 2808  
2761 CGACTCGTGTGTCGACCGGTACCGCTCCGGGATCGGTCAAGTGGCCACATCCAAAC 2820  
2809 ATGAGGAGTGGTCTGTGACGACCTGGAGAAATCCCTTTTATGGCAAGGATCCCTCC 2868  
2821 ATGAGGAGTGGTCTGTGACGACCTGGAGAAATCCCTTTTATGGCAAGGATCCCTCC 2880  
2869 ATCGAGACCATCAAGGGGGGAGGCCTCATTTTCTGCAATCCCAAGAGAAATGTAT 2928  
2881 ATCGAGACCATCAAGGGGGGAGGCCTCATTTTCTGCAATCCCAAGAGAAATGTAT 2940  
2929 GAGTCCCGCGAAGCTGTCCGGGCTCGGACTCAATGCTGATGATATATACCGGGGCTT 2988  
2941 GAGTCCCGCGAAGCTGTCCGGGCTCGGACTCAATGCTGATGATATATACCGGGGCTT 3000

2989 GATGTATCGGTATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3048  
3001 GATGTATCGGTATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3060  
3049 ACGGGCTTTACCGGGGATTTGACTCAGTGAATCGACTGCAATACATGTGTCAACAGACA 3108  
3061 ACGGGCTTTACCGGGGATTTGACTCAGTGAATCGACTGCAATACATGTGTCAACAGACA 3120  
3109 GTGCACTTCAGCTGGAACCGGACCTTACCATTTAGAGACGACCGTGCACAAAGCGG 3168  
3121 GTGCACTTCAGCTGGAACCGGACCTTACCATTTAGAGACGACCGTGCACAAAGCGG 3180  
3169 GTGTACAGCTCGCAGCGGAGGAGGACTGTGTAGGGGAGATGGGCAATTTACAGTTT 3228  
3181 GTGTACAGCTCGCAGCGGAGGAGGACTGTGTAGGGGAGATGGGCAATTTACAGTTT 3240  
3229 GTGACTTCCAGGAGAACCGGCCCTCGGGCATGTTTCGATTTCTCGGTCTGTGCGAGTCTAT 3288  
3241 GTGACTTCCAGGAGAACCGGCCCTCGGGCATGTTTCGATTTCTCGGTCTGTGCGAGTCTAT 3300  
3289 GAGCGGGCTGTGCTGTGACTACGAGCTCAGCCCCCGGAGACCTCAGTTAGTTGCGGCT 3348  
3301 GAGCGGGCTGTGCTGTGACTACGAGCTCAGCCCCCGGAGACCTCAGTTAGTTGCGGCT 3360  
3349 TACCTAAACACACCGAGGTTGCCGCTCTGCCAGGACATCTGGAGTTCTGGGAGAGCGTC 3408  
3361 TACCTAAACACACCGAGGTTGCCGCTCTGCCAGGACATCTGGAGTTCTGGGAGAGCGTC 3420  
3409 TTTTACAGGCTTACCCACATAGACGCCCATTTCTTCTCCAGACTAAGCAGGAGGAGAC 3468  
3421 TTTTACAGGCTTACCCACATAGACGCCCATTTCTTCTCCAGACTAAGCAGGAGGAGAC 3480  
3469 AACTTCCCTTACTGTGTAGTACAGGCTACGGTGTGGCGGAGGCTCAGGCTCCACCT 3528  
3481 AACTTCCCTTACTGTGTAGTACAGGCTACGGTGTGGCGGAGGCTCAGGCTCCACCT 3540  
3529 CCATCTGGGACCAATGTGGAAGTCTCTATACGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 3588  
3541 CCATCTGGGACCAATGTGGAAGTCTCTATACGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 3600  
3589 ACCCCCTCTGTGTATAGGCTGGGAGCGGTTTCAAACGAGGTTTAACTACCAACACACCCCAT 3648  
3601 ACCCCCTCTGTGTATAGGCTGGGAGCGGTTTCAAACGAGGTTTAACTACCAACACACCCCAT 3660  
3649 ACCAAATACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3708  
3661 ACCAAATACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720  
3709 CTGTTAGGGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGTGACAAACAGGAGCGTGGTC 3768  
3721 CTGTTAGGGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGTGACAAACAGGAGCGTGGTC 3780  
3769 ATTGTGGGAGGATCATTTTGTCCGAAAGCGGCCCATCATTTCCCGACAGGGAAGTCCCT 3828  
3781 ATTGTGGGAGGATCATTTTGTCCGAAAGCGGCCCATCATTTCCCGACAGGGAAGTCCCT 3840  
3829 TACCGGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3888  
3841 TACCGGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900  
3889 ATGAGCTCCCGCAACAATTTCAAACAGAGGCAATCGGGTTGCTGCAAAACAGCCACCAAG 3948  
3901 ATGAGCTCCCGCAACAATTTCAAACAGAGGCAATCGGGTTGCTGCAAAACAGCCACCAAG 3960  
3949 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4008  
3961 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4020  
4009 GCGAAGCATATGTGAATTTTCATCAGCGGATNATATATTTAGAGGCTTGTCCACTCTG 4068  
4021 GCGAAGCATATGTGAATTTTCATCAGCGGATNATATATTTAGAGGCTTGTCCACTCTG 4080  
4069 CTTGGCAACCCCGGATAGCATCACTGATGGCATTCAGGCTTATCAGGCTTATCAGGCGGCTC 4128

Db 4081 ||||| CCTGGCAACCCCGGATAGCATCACTGATGGCAATTCAGAGCTCTATACACGCGCTC 4140  
QY 4129 ACCACCCACATACCTCTGTTTAAATCTCTGGGGGATGGTGGCCGCCAACTTGCT 4188  
Db 4141 ACCACCAACATACCTCTGTTTAAATCTCTGGGGGATGGTGGCCGCCAACTTGCT 4200  
QY 4189 CCTCCAGCGCTGCTCTCTCTTCTGTTAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC 4248  
Db 4201 CCTCCAGCGCTGCTCTCTCTTCTGTTAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC 4260  
QY 4249 ATAGGCTTGGGAAGGTGCTGTGGATATTTGGCAGGTTATGGACAGGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTGGGAAGGTGCTGTGGATATTTGGCAGGTTATGGACAGGGGTGGCAGGC 4320  
QY 4309 GCGCTCGTGGCTTTAAAGTCATGACGCGGAGATGCCCTCACCGAGACCTGTTAAC 4368  
Db 4321 GCGCTCGTGGCTTTAAAGTCATGACGCGGAGATGCCCTCACCGAGACCTGTTAAC 4380  
QY 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTGGTGGTGGCGCAGGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTGGTGGTGGCGCAGGATA 4440  
QY 4429 CTGCGTGGGCACTGGGCGCCAGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGC 4488  
Db 4441 CTGCGTGGGCACTGGGCGCCAGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGC 4500  
QY 4489 TTCGCTTCGGGGTAAACAGCTCTCCCGCACGACTATGCTGAGAGCGACGCTGA 4548  
Db 4501 TTCGCTTCGGGGTAAACAGCTCTCCCGCACGACTATGCTGAGAGCGACGCTGA 4560  
QY 4549 GCACGTGCTCACTCAGATCCTCTCTAGTCTTACCATCACTCACTGCTGAAGGCTTCAC 4608  
Db 4561 GCACGTGCTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTCTGAGAGGCTTCAC 4620  
QY 4609 CAGTGGATCAAGGAGATGTGCCAGGCACTCCGGCTCGTGGCTGAAGATGTTGG 4668  
Db 4621 CAGTGGATCAAGGAGATGTGCCAGGCACTCCGGCTCGTGGCTGAAGATGTTGG 4680  
QY 4669 GATTGGATGACAGGCTGCTGACTGATTTCAAGCTTGCTCCAGTCCAGTCTCTGCGC 4728  
Db 4681 GATTGGATGACAGGCTGCTGACTGATTTCAAGCTTGCTCCAGTCCAGTCTCTGCGC 4740  
QY 4729 CGATTGCGGGAGTCCCTCTCTCTCATGTCAACGTTGGGTAACAGGAGTCTGGCGGGC 4788  
Db 4741 CGATTGCGGGAGTCCCTCTCTCTCATGTCAACGTTGGGTAACAGGAGTCTGGCGGGC 4800  
QY 4789 GACGGCATATGCAAAACCACTGCGCCATGTGGAGCACAGATCAACCGGACATGTGAAAAC 4848  
Db 4801 GACGGCATATGCAAAACCACTGCGCCATGTGGAGCACAGATCAACCGGACATGTGAAAAC 4860  
QY 4849 GTTTCATGAGATCGTGGGCTAGGACCTGTAGTAAACAGTGGATGACATGGACATTCGCC 4908  
Db 4861 GTTTCATGAGATCGTGGGCTAGGACCTGTAGTAAACAGTGGATGACATGGACATTCGCC 4920  
QY 4909 ATTAACCGGTACACCAAGGCGCCCTGACGCGCCCTCCCGGGGCCAAATATTCTAGGGCG 4968  
Db 4921 ATTAACCGGTACACCAAGGCGCCCTGACGCGCCCTCCCGGGGCCAAATATTCTAGGGCG 4980  
QY 4969 CTGTGGGGGTGGCTGTGAGAGTACCTGAGAGTACCGGGGTGGGGAATTCACATAC 5028  
Db 4981 CTGTGGGGGTGGCTGTGAGAGTACCTGAGAGTACCGGGGTGGGGAATTCACATAC 5040  
QY 5029 GTGACGGGCTATGACCACTGACAACTGTAAGTGGCGGTGCTGAGTTCGGGCGCCGAAATTC 5088  
Db 5041 GTGACGGGCTATGACCACTGACAACTGTAAGTGGCGGTGCTGAGTTCGGGCGCCGAAATTC 5100  
QY 5089 TTCAAGAGTGGATGGGCTGCGGTGCAAGGTAACGCTCCAGCGTGAACACCCCTCTTA 5148  
Db 5101 TTCAAGAGTGGATGGGCTGCGGTGCAAGGTAACGCTCCAGCGTGAACACCCCTCTTA 5160  
QY 5149 CGGAGGAGTCACTTCTGGTGGGCTCAATCAATACCTGGTGGGTCAAGCTCCCA 5208

Db 5161 CGGAGGAGGTCACTTCTGCTGGGCTCAATCAATCACTGTTGGGTCAAGCTCCCA 5220  
QY 5209 TCGAGCCCGAACCAGACGTAGAGTGTCACTTCATGCTCAACGACCCCTCCCACTT 5268  
Db 5221 TCGAGCCCGAACCAGACGTAGAGTGTCACTTCATGCTCAACGACCCCTCCCACTT 5280  
QY 5269 AGCGGGAGACGGCTAAGCGTAGGTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCA 5328  
Db 5281 AGCGGGAGACGGCTAAGCGTAGGTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGCTAGCAGCTGTCTGCGCTTCTTTGAAGGAAATGCACTACCCGTCACTACTCC 5388  
Db 5341 TCAGCTAGCAGCTGTCTGCGCTTCTTTGAAGGAAATGCACTACCCGTCACTACTCC 5400  
QY 5389 CGGACGCTGACCTCATGAGGCCAACTCTCTGTGGCGGACGAGATGGCGGGAACTC 5448  
Db 5401 CGGACGCTGACCTCATGAGGCCAACTCTCTGTGGCGGACGAGATGGCGGGAACTC 5460  
QY 5449 ACCCGCTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTTCCAAAGC 5508  
Db 5461 ACCCGCTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTTCCAAAGC 5520  
QY 5509 GAGGAGGATGAGAGGAAAGTATCCCTTCGCGGAGATCTCTCGGAGTCCAGGAAATTC 5568  
Db 5521 GAGGAGGATGAGAGGAAAGTATCCCTTCGCGGAGATCTCTCGGAGTCCAGGAAATTC 5580  
QY 5569 CCTGAGCGATGCCCATATGAGGACGCGCGGANTTAAACCCCTTCACTGTTAGTCTCTGG 5628  
Db 5581 CCTGAGCGATGCCCATATGAGGACGCGCGGANTTAAACCCCTTCACTGTTAGTCTCTGG 5640  
QY 5629 AAGGACCCGGAATAGTCTCCCTCAGTGTGACAGGCTGTCATTCGCGCTCCAGGCGC 5688  
Db 5641 AAGGACCCGGAATAGTCTCCCTCAGTGTGACAGGCTGTCATTCGCGCTCCAGGCGC 5700  
QY 5689 CCTCGATACCACTTCCACGGAGGAGGAGCGTGTCTGTCTGATCAGAAATCTACCGTGTCT 5748  
Db 5701 CCTCGATACCACTTCCACGGAGGAGGAGCGTGTCTGTCTGATCAGAAATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGCGGAGCTCGCCAAAGACCTTTCGCGAGTCCGAATCTCGCGCTCGAC 5808  
Db 5761 TCTGCTTGGCGGAGCTCGCCAAAGACCTTTCGCGAGTCCGAATCTCGCGCTCGAC 5820  
QY 5809 AGCGGACCGCAACGGGCTCTCTGACGAGCCCTCCGACGAGCGCGAGCGGATCCGAC 5868  
Db 5821 AGCGGACCGCAACGGGCTCTCTGACGAGCCCTCCGACGAGCGCGAGTCCGAC 5880  
QY 5869 GTTGAGTCTACTCTCTCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTACTCTCTCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGTCTTGGTCTACCGTAAAGCGAGGAGCTAGTAGGAGCTGTCTGTCTCGATG 5988  
Db 5941 GACGGTCTTGGTCTACCGTAAAGCGAGGAGCTAGTAGGAGCTGTCTGTCTCGATG 6000  
QY 5989 TCCTACATGACGAGCGCTGATACGCGATGCGCTCGCGGAGGAAACCAAGTCCG 6048  
Db 6001 TCCTACATGACGAGCGCTGATACGCGATGCGCTCGCGGAGGAAACCAAGTCCG 6060  
QY 6049 ATCAATGACGTGAGCACTCTTGTCTCGTCAACCAACTTGGTCTATGTCTACACATCT 6108  
Db 6061 ATCAATGACGTGAGCACTCTTGTCTCGTCAACCAACTTGGTCTATGTCTACACATCT 6120  
QY 6109 CGGAGCGCAAGCTTGGCGGAGAGGATCACTTTGACAGCTGAGGCTCTTGACGAC 6168  
Db 6121 CGGAGCGCAAGCTTGGCGGAGAGGATCACTTTGACAGCTGAGGCTCTTGACGAC 6180  
QY 6169 CACTACCGGAGCTGTCTCAAGGAGATGAAGCGGAGGCTCCACAGTCTAAGCTTAACTT 6228  
Db 6181 CACTACCGGAGCTGTCTCAAGGAGATGAAGCGGAGGCTCCACAGTCTAAGCTTAACTT 6240  
QY 6229 CTATCGTGGAGGAGCTGTAAAGCTGACGCCCCCAATTCGGCGAGATCTAAATTTGGC 6288  
Db 6241 CTATCGTGGAGGAGCTGTAAAGCTGACGCCCCCAATTCGGCGAGATCTAAATTTGGC 6300

QY	6289	TATGGGCAAGGACGTCGGAAACCTATCCAGCNAAGCCGTTAAACACATCCGCTCGGTG	6348
DB	6301	TATGGGCAAGGACGTCGGAAACCTATCCAGCAAGCCGTTAAACACATCCGCTCGGTG	6360
QY	6349	TGGAAGGACATTGCTGGAAGACACTGACAGACACCAANTTGACACCAACCATCATGGCAAAAAT	6408
DB	6361	TGGNAGACCTTGCTGGAAGACACTGAGACACCAANTTGACACCAACCATCATGGCAAAAAT	6420
QY	6409	GAGGTTTTCTGCGTCCAACACAGAAAGGGGGCCGCAAGCAGCTCGCCCTTATCGTATTTC	6468
DB	6421	GAGGTTTTCTGCGTCCAACACAGAAAGGGGGCCGCAAGCAGCTCGCCCTTATCGTATTTC	6480
QY	6469	CCAGATTTGGGGGTTCTGTGTGTCGGAGAAAATGGCCCTTTACGATGPGTCTCCACCCCTC	6528
DB	6481	CCAGATTTGGGGGTTCTGTGTGTCGGAGAAAATGGCCCTTTACGATGPGTCTCCACCCCTC	6540
QY	6529	CCTCAGCCCTGATGGGCTCTTCCATCGGAATCCAAATACCTCTCTGACAGCGGGTCGAG	6588
DB	6541	CCTCAGCCCTGATGGGCTCTTCCATCGGAATCCAAATACCTCTCTGACAGCGGGTCGAG	6600
QY	6589	TTCTGTGTGAATGCTCGTGAAGCGAAATAATGCCCTATGGGCTTCGCATATGACACCCGC	6648
DB	6601	TTCTGTGTGAATGCTCGTGAAGCGAAATAATGCCCTATGGGCTTCGCATATGACACCCGC	6660
QY	6649	TGTTTGTACTCAAGGFTCACTGAGATGACATCGTGTTGAGGAGTCAATCTACCAATGT	6708
DB	6661	TGTTTGTACTCAAGGFTCACTGAGATGACATCGTGTTGAGGAGTCAATCTACCAATGT	6720
QY	6709	TGTGACTTTGGCCCGGAAAGCAGACAGCCCATAGGTCGCTCAGAGCGGCTTTACATC	6768
DB	6721	TGTGACTTTGGCCCGGAAAGCAGACAGCCCATAGGTCGCTCAGAGCGGCTTTACATC	6780
QY	6769	GGGGCCCCCTGACTAATTTAAAGGCGCAAACTGCGGCTATCGCCGTCGCCGCGAGC	6828
DB	6781	GGGGCCCCCTGACTAATTTAAAGGCGCAAACTGCGGCTATCGCCGTCGCCGCGAGC	6840
QY	6829	GGTGTACTGACGACAGCTCGGCTAATPACCTCACATGTTACTTGAAGCGCGCTGCGGC	6888
DB	6841	GGTGTACTGACGACAGCTCGGCTAATPACCTCACATGTTACTTGAAGCGCGCTGCGGC	6900
QY	6889	TGTCGAGCTCGGAAGCTCAGGACTGCAGATGCTCGTATGCGGAGACGCTTGTGCTT	6948
DB	6901	TGTGAGCTCGGAAGCTCAGGACTGCAGATGCTCGTATGCGGAGACGCTTGTGCTT	6960
QY	6949	ATCTGTAAAGCGGGGACCCAAAGAGCAGAGCGAGCGCTACGGGCTTTCAGCGAGGCT	7008
DB	6961	ATCTGTAAAGCGGGGACCCAAAGAGCAGAGCGAGCGCTACGGGCTTTCAGCGAGGCT	7020
QY	7009	ATGACTAGATPACTCTGCCCCCTTGGGGACCCGCCCAAACAGAAATACGATTTGGAGTTG	7068
DB	7021	ATGACTAGATPACTCTGCCCCCTTGGGGACCCGCCCAAACAGAAATACGATTTGGAGTTG	7080
QY	7069	ATAACATCATGCTCCTCAATGTGTAGTCGCGCAGCATGCTGCGCAAAAGGTTGTAC	7128
DB	7081	ATAACATCATGCTCCTCAATGTGTAGTCGCGCAGCATGCTGCGCAAAAGGTTGTAC	7140
QY	7129	TATCTACCCGTCAGCCACACCCCTTGGCGGCTGCGGAGCAGGTGAGAGCAGCTAGACAC	7188
DB	7141	TATCTACCCGTCAGCCACACCCCTTGGCGGCTGCGGAGCAGGTGAGAGCAGCTAGACAC	7200
QY	7189	ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGATGCGGCCACTTTGTGGGCAAGG	7248
DB	7201	ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGATGCGGCCACTTTGTGGGCAAGG	7260
QY	7249	ATGATCTGATGACTCATTTCTCTCATGCTTAGCTCAGGAAACAACTTGAAAAAGCC	7308
DB	7261	ATGATCTGATGACTCATTTCTCTCCATCTTCTAGCTCAGGAAACAACTTGAAAAAGCC	7320
QY	7309	CTGATTTGTCAGATCTACGGGCTGTTACTCCATTGAGCACTTGACCTACCTCAGATC	7368
DB	7321	CTGATTTGTCAGATCTACGGGCTGTTACTCCATTGAGCACTTGACCTACCTCAGATC	7380

### RESULT 3

```

US-09-539-601-22
; Sequence 22, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 8001
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/NS3-3'/5.1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:

```

```

; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3 - 5B
; OTHER INFORMATION: of cell culture-adapted clone no. 5.1
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (7771)..(8001)
; US-09-539-601-22

Query Match          99.5%; Score 7947.8; DB 4; Length 8001;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 7977; Conservative 0; Mismatches 12; Indels 12; Gaps 1;

QY 1 GCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Db 1 GCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60

QY 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGCAAGCTTCCAGGAC 120
Db 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGCAAGCTTCCAGGAC 120

QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAACCAGGTGAGTACACCGGAATTCGCCAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAACCAGGTGAGTACACCGGAATTCGCCAG 180

QY 181 GACGACCGGGTCTTTCTTGGATCAACCCCGCTCAATGCCCTGGAGATTGGGGCGTGCCCC 240
Db 181 GACGACCGGGTCTTTCTTGGATCAACCCCGCTCAATGCCCTGGAGATTGGGGCGTGCCCC 240

QY 241 GCGAGACTCTAGCCAGTAGTCTTGGTTCGCGAAGAGGCTTGTGCTACTGCTGATAG 300
Db 241 GCGAGACTCTAGCCAGTAGTCTTGGTTCGCGAAGAGGCTTGTGCTACTGCTGATAG 300

QY 301 GTGCTTGCAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGACGACCAATTCCTAAAC 360
Db 301 GTGCTTGCAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGACGACCAATTCCTAAAC 360

QY 361 CTCAAGAAAAAACCCAAA-----GGGCGCGCCATGATTGAACAAGATTGGATTGC 408
Db 361 CTCAAGAAAAAACCCAAA-----GGGCGCGCCATGATTGAACAAGATTGGATTGC 408

QY 409 AGCAGGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACACAGA 468
Db 409 AGCAGGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACACAGA 468

QY 421 AGCAGGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACACAGA 480
Db 421 AGCAGGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACACAGA 480

QY 469 CAATCGGCTGCTCTGATCGCGCGTTCGGGCTGTGACGCGGAGGGGCGCGGTTCTTT 528
Db 469 CAATCGGCTGCTCTGATCGCGCGTTCGGGCTGTGACGCGGAGGGGCGCGGTTCTTT 528

QY 481 CAATCGGCTGCTCTGATCGCGCGTTCGGGCTGTGACGCGGAGGGGCGCGGTTCTTT 540
Db 481 CAATCGGCTGCTCTGATCGCGCGTTCGGGCTGTGACGCGGAGGGGCGCGGTTCTTT 540

QY 529 TTGTCAAGACCGACCTGTCCGTCGCTTGAATGAATGACGAGGACGAGGCGCGCTAT 588
Db 529 TTGTCAAGACCGACCTGTCCGTCGCTTGAATGAATGACGAGGACGAGGCGCGCTAT 588

QY 541 TTGTCAAGACCGACCTGTCCGTCGCTTGAATGAATGACGAGGACGAGGCGCGCTAT 600
Db 541 TTGTCAAGACCGACCTGTCCGTCGCTTGAATGAATGACGAGGACGAGGCGCGCTAT 600

QY 589 CGTGGCTGCGACGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 648
Db 589 CGTGGCTGCGACGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 648

QY 601 CGTGGCTGCGACGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 660
Db 601 CGTGGCTGCGACGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 660

QY 649 GAAGGACTGGGCTGCTATTGGGGGAGAGTGCAGGGGAGGATCTCCTGTCACTCACTTG 708
Db 649 GAAGGACTGGGCTGCTATTGGGGGAGAGTGCAGGGGAGGATCTCCTGTCACTCACTTG 708

QY 661 GAAGGACTGGGCTGCTATTGGGGGAGAGTGCAGGGGAGGATCTCCTGTCACTCACTTG 720
Db 661 GAAGGACTGGGCTGCTATTGGGGGAGAGTGCAGGGGAGGATCTCCTGTCACTCACTTG 720

QY 709 CTCTTCGAGAAAGATTCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 768
Db 709 CTCTTCGAGAAAGATTCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 768

QY 721 CTCTTCGAGAAAGATTCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 780
Db 721 CTCTTCGAGAAAGATTCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 780

QY 769 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGGACGCTACTTCGGA 828
Db 769 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGGACGCTACTTCGGA 828

QY 781 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGGACGCTACTTCGGA 840
Db 781 CGGCTACCTGCCATTCGACCAACCAAGCAACATCGCATCGAGCGGACGCTACTTCGGA 840
```

```

QY 829 TGGAAAGCCGCTCTTGTCCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCCAG 888
Db 841 TGGAAAGCCGCTCTTGTCCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCCAG 900

QY 889 CCGAACTGTTCCCGAGGCTCAAGGGCGCATGCCGACGGGAGGATCTGTCGTGACCC 948
Db 901 CCGAACTGTTCCCGAGGCTCAAGGGCGCATGCCGACGGGAGGATCTGTCGTGACCC 960

QY 949 ATGGCGATGCTGCTTGGCCGAATATCATGTTGAAATGCGCGCTTCTTGATTCATCG 1008
Db 961 ATGGCGATGCTGCTTGGCCGAATATCATGTTGAAATGCGCGCTTCTTGATTCATCG 1020

QY 1009 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGGTTGGTACCCGTGATA 1068
Db 1021 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGGTTGGTACCCGTGATA 1080

QY 1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACGGCTTCTGTCGTTTACGTTATCGCG 1128
Db 1081 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACGGCTTCTGTCGTTTACGTTATCGCG 1140

QY 1129 CTCGCGATTTCGAGGSCATCGGCTTCTATCGGCTTCTTACGAGTTCTTCTGAGTTTAA 1188
Db 1141 CTCGCGATTTCGAGGSCATCGGCTTCTATCGGCTTCTTACGAGTTCTTCTGAGTTTAA 1200

QY 1189 CAGACCAACACGTTTCCCTCTAGCGGGAATCAATTCCGCGCCCTCTCCCTCCCCCCCC 1248
Db 1201 CAGACCAACACGTTTCCCTCTAGCGGGAATCAATTCCGCGCCCTCTCCCTCCCCCCCC 1260

QY 1249 AAGCTTACTGGCGGAAGCGCTTGGAAATAGGCGGCTGCGTTGCTATATGTTATTT 1308
Db 1261 AAGCTTACTGGCGGAAGCGCTTGGAAATAGGCGGCTGCGTTGCTATATGTTATTT 1320

QY 1309 TCCACCATATCCGCTCTTTTGGCAATGAGGCGCGGAAACCTGCGCCCTGCTCTCTTG 1368
Db 1321 TCCACCATATCCGCTCTTTTGGCAATGAGGCGCGGAAACCTGCGCCCTGCTCTCTTG 1380

QY 1369 ACGAGCATTCCTAGGGGCTTTTCCCTCTCGCAAGAGGAATCAAGGTCGTGTAATGTC 1428
Db 1381 ACGAGCATTCCTAGGGGCTTTTCCCTCTCGCAAGAGGAATCAAGGTCGTGTAATGTC 1440

QY 1429 GTGAAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGCAACAAACAGTCTGTAGGACCTT 1488
Db 1441 GTGAAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGCAACAAACAGTCTGTAGGACCTT 1500

QY 1489 TCGAGGACGAGGAAACCCCGCTGCGGACAGTGTCTCTGCGGCAAGAACACACGTA 1548
Db 1501 TCGAGGACGAGGAAACCCCGCTGCGGACAGTGTCTCTGCGGCAAGAACACACGTA 1560

QY 1549 TAAGATACACCTGCAAGGCGGCAACCCCGCTGCGGACAGTGTGTAAGTGTGTG 1508
Db 1561 TAAGATACACCTGCAAGGCGGCAACCCCGCTGCGGACAGTGTGTAAGTGTGTG 1620

QY 1609 GAAAGATCAATGGCTCTCTCAAGCGTATCAACAGGGGCTGAAGATGCCAGAAG 1668
Db 1621 GAAAGATCAATGGCTCTCTCAAGCGTATCAACAGGGGCTGAAGATGCCAGAAG 1680

QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGACATGCTTTACATGTTTAG 1728
Db 1681 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGACATGCTTTACATGTTTAG 1740

QY 1729 TCGAGTTAAAAAACCGCTTAGGCCCCCGGAAACACCGGGGACGTGGTTTTCTTTGAAAA 1788
Db 1741 TCGAGTTAAAAAACCGCTTAGGCCCCCGGAAACACCGGGGACGTGGTTTTCTTTGAAAA 1800

QY 1789 CAGGATATACCATGGCGCTTATAGGCTTACTCCAAACAGACGCGGACCTACTTGGC 1848
Db 1801 CAGGATATACCATGGCGCTTATAGGCTTACTCCAAACAGACGCGGACCTACTTGGC 1860

QY 1849 TGCATCATCACTAGCTTACAGCGGAGACAGGAAACAGTCCGAGGGGAGGTCCAAAGT 1908
Db 1861 TGCATCATCACTAGCTTACAGCGGAGACAGGAAACAGTCCGAGGGGAGGTCCAAAGT 1920
```

1909 GTCTCCACCGCAACACAACTTTCTCTGCGACCTCGCTCAATGCGGTGCTGTGACTGTC 1968  
1921 GTCTCCACCGCAACAACTTTCTCTGCGACCTCGCTCAATGCGGTGCTGTGACTGTC 1980  
1969 TATCATGGTGC CGGCTCAAGACCTTTGCGGCGCCAAAGGCGCCCAATCACCACCAATGTAC 2028  
1981 TATCATGGTGC CGGCTCAAGACCTTTGCGGCGCCAAAGGCGCCCAATCACCACCAATGTAC 2040  
2029 ACCAATGTGACACGAGGACTGCTGGCTGGCAAGGCGCCCGCGGGCGCGTTCCTTGACA 2088  
2041 ACCAATGTGACACGAGGACTGCTGGCTGGCAAGGCGCCCGCGGGCGCGTTCCTTGACA 2100  
2089 CCATGCACTTGC CGGCTCGGACCTTTACTTGTGTACAGAGCATGCCGATGCTCAATCCG 2148  
2101 CCATGCACTTGC CGGCTCGGACCTTTACTTGTGTACAGAGCATGCCGATGCTCAATCCG 2160  
2149 GTGCGCGCGGCGGCGAGCAGCAGGGGAGCCTACTCTCCCCCAGGCGCGCTCTCCTACTTG 2208  
2161 GTGCGCGCGGCGGCGAGCAGCAGGGGAGCCTACTCTCCCCCAGGCGCGCTCTCCTACTTG 2220  
2209 AAGGGCTTTGCGGCGGTTCACCTGCTGCGGCTCGGCGCACGCTGTGGGCACTTTTCGG 2268  
2221 AAGGGCTTTGCGGCGGTTCACCTGCTGCGGCTCGGCGCACGCTGTGGGCACTTTTCGG 2280  
2269 GCTGCCGTGTGCACCCGAGGGGTTTGCAGAGCGGTGGACTTTGTACCGCTCGAGTCTATG 2328  
2281 GCTGCCGTGTGCACCCGAGGGTTTGCAGAGCGGTGGACTTTGTACCGCTCGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCCGGTCTTCAAGCAACTCGTCTCCCCCGGCGGTACCGCAG 2388  
2341 GAAACCACTATGCGGTCCCCGGTCTTCAAGCAACTCGTCTCCCCCGGCGGTACCGCAG 2400  
2389 ACATTCAGGTGCGCATCTACAGCGCCCTACTGGTAGCGCAAGCAGCTAAGGTGCGG 2448  
2401 ACATTCAGGTGCGCATCTACAGCGCCCTACTGGTAGCGCAAGCAGCTAAGGTGCGG 2460  
2449 GCTGCGTATGCAAGGCTTAAGGTGCTGCTCTGAACCGCTCGCTGCGCGCCAC 2508  
2461 GCTGCGTATGCAAGGCTTAAGGTGCTGCTGAACCGCTCGCTGCGCGCCAC 2520  
2509 CTAGTTTCGGGCGGTATATGCTPAAGGCACATGGTATCGACCTTAACATCAGAACCGGG 2568  
2521 CTAGTTTCGGGCGGTATATGCTPAAGGCACATGGTATCGACCTTAACATCAGAACCGGG 2580  
2569 GTAAGACCACTACACCGGTGCCCCCATCAGTACTCCAGCTATGGAAGTTTCTTGCC 2628  
2581 GTAAGACCACTACACCGGTGCCCCCATCAGTACTCCAGCTATGGAAGTTTCTTGCC 2640  
2629 GACGCTGCTTGTCTGCGGGCGCTATGACATCATAATATGTATGATGATGCTCACTCACT 2688  
2641 GACGCTGCTTGTCTGCGGGCGCTATGACATCATAATATGTATGATGATGCTCACTCACT 2700  
2689 GACTCGACCACTATCTTGGGCGATCGGCACAGTCTCTGACCAAGCGAGAGCGGTGGAGCG 2748  
2701 GACTCGACCACTATCTTGGGCGATCGGCACAGTCTCTGACCAAGCGAGAGCGGTGGAGCG 2760  
2749 CGACTCGCTGCTCGCCACCGCTACCGCTCGGGATCGGTACCGTGCACATFCCAAAC 2808  
2761 CGACTCGCTGCTGCTCGCCACCGCTACCGCTCGGGATCGGTACCGTGCACATFCCAAAC 2820  
2809 ATCGAGAGGTGGTCTGTCTGACGACCTGGAGAAATCCCTTTTATGCAAGCCATCCCC 2868  
2821 ATCGAGAGGTGGTCTGTCTGACGACCTGGAGAAATCCCTTTTATGCAAGCCATCCCC 2880  
2869 ATCGAGACCACTAAGGGGGGAGGACCTCATTTTCTGCCATTCCTAAGAGAAATGTGAT 2928  
2881 ATCGAGACCACTAAGGGGGGAGGACCTCATTTTCTGCCATTCCTAAGAGAAATGTGAT 2940  
2929 GAGCTCGCGGAGAGCTGTGCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGCGCTT 2988  
2941 GAGCTCGCGGAGAGCTGTGCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGCGCTT 3000  
2989 GATGTATCCGTTCATAACCACTAGCGGAGACGTCTATTTGCTGTAGCAACGAGCGCTCTAATG 3048

3001 GATGTATCCGTTCATACCACTAGCGGAGACGTCTATTTGCTGTGACCAACGAGCGCTCTAATG 3060  
3049 ACGGGCTTTACCGCGGATTTTGCAGTCACTGATCACTGCAATCATATGTGTCAACAGACA 3108  
3061 ACGGGCTTTACCGCGTGTGCTTGCAGTCACTGATCACTGCAATCATATGTGTCAACAGACA 3120  
3109 GTGCACTTACGCTTGAGACCGGACCTTCACTTACAGACGAGCGGTGCCACAGACGCG 3168  
3121 GTGCACTTACGCTTGAGACCGGACCTTCACTTACAGACGAGCGGTGCCACAGACGCG 3180  
3169 GTGTCAGCTCGCAGCGCGGAGGAGGAGTGTAGGCGGAGGATGGGCAATTTACAGGTTT 3228  
3181 GTGTCAGCTCGCAGCGCGGAGGAGGAGTGTAGGCGGAGGATGGGCAATTTACAGGTTT 3240  
3229 GTGCACTCCAGGAGAACCGGCGCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTAT 3288  
3241 GTGACTCCAGGAGAACCGGCGCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTAT 3300  
3289 GACCGGGCTGTGCTTGTGTACGAGCTCACGCGCGCGGAGACCTCAGTTAGTTGCGGGCT 3348  
3301 GACCGGGCTGTGCTTGTGTACGAGCTCACGCGCGCGGAGACCTCAGTTAGTTGCGGGCT 3360  
3349 TACCTAAACACACACAGGAGTTGCCGCTCTGCCAGAACCATCTGGAGTTCTGGGAGAGGTC 3408  
3361 TACCTAAACACACACAGGAGTTGCCGCTCTGCCAGAACCATCTGGAGTTCTGGGAGAGGTC 3420  
3409 TTTACAGGCGCTCACCCACATAGAGCGCCATTTCTTTGCCAGACTAAGCAGGAGGAGAC 3468  
3421 TTTACAGGCGCTCACCCACATAGAGCGCCATTTCTTTGCCAGACTAAGCAGGAGGAGAC 3480  
3469 AACTTCCCTTACCTGTGTAGCATACGAGCTACGCTGTGCGGCGAGGCTCAGGCTCCACT 3528  
3481 AACTTCCCTTACCTGTGTAGCATACGAGCTACGCTGTGCGGCGAGGCTCAGGCTCCACT 3540  
3529 CCATCTGGGACCAAAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCA 3588  
3541 CCATCTGGGACCAAAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCA 3600  
3589 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACACCCATA 3648  
3601 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACACCCATA 3660  
3649 ACCAATACATCATGGCATGCTGCGCTGACCTGTGAGTGTGTCACAGACCTTGGGTG 3708  
3661 ACCAATACATCATGGCATGCTGCGCTGACCTGTGAGTGTGTCACAGACCTTGGGTG 3720  
3709 CTGCTAGGCGGAGTCTTACGAGCTCTGCGCGGCTGCTGCAACAGCAGCGGTGCTC 3768  
3721 CTGCTAGGCGGAGTCTTACGAGCTCTGCGCGGCTGCTGCAACAGCAGCGGTGCTC 3780  
3769 ATTGTGGCGAGGATCATTTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCTCTT 3828  
3781 ATTGTGGCGAGGATCATTTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCTCTT 3840  
3829 TACCGGAGTTTCATGAGATGGAAGTGTGCGCTCACACCTCCCTTTACATCGAACAGGGA 3888  
3841 TACCGGAGTTTCATGAGATGGAAGTGTGCGCTCACACCTCCCTTTACATCGAACAGGGA 3900  
3889 ATCAGCTCGCGCAACAAATTCAAACAGAGGCAATCGGTTGCTGCAACAGCAGCCCAAG 3948  
3901 ATCAGCTCGCGCAACAAATTCAAACAGAGGCAATCGGTTGCTGCAACAGCAGCCCAAG 3960  
3949 CAAGCGAGGCTGCTGCTCCCGTGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4008  
3961 CAAGCGAGGCTGCTGCTCCCGTGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4020  
4009 GCGAAGCATATGTGGAATTTTCATCAGGGGATACAAATTTTAGAGGCTTGTCCACTCTG 4068  
4021 GCGAAGCATATGTGGAATTTTCATCAGGGGATACAAATTTTAGAGGCTTGTCCACTCTG 4080  
4069 CTGTGGCAACCCCGGATAGCATCACTGATGGCAATTCACAGCTCTATCACCGCCGCTC 4128

Db 4081 CCTGGCAACCCCGGATAGCATCACTGATGGCAATTCACAGCCTCTATCACCAGCCCGCTC 4140  
Qy 4129 ACCACCAACATACCTCTCTGTTTAAATCATCTCTGGGGGATGGTGGCCGCCCACTTGCT 4188  
Db 4141 ACCACCAACATACCTCTCTGTTTAAATCATCTCTGGGGGATGGTGGCCGCCCACTTGCT 4200  
Qy 4189 CTTCCAGCGCTGCTCTCTCTCTGTTGFPAGSGCCCGGCATCGCTGAGCGGCTGTTGCGAGC 4248  
Db 4201 CTTCCAGCGCTGCTCTCTCTCTGTTGFPAGSGCCCGGCATCGCTGAGCGGCTGTTGCGAGC 4260  
Qy 4249 ATAGGCTTTGGGAAGTGTCTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTTGGGAAGTGTCTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4320  
Qy 4309 GGGCTGTCGGCTTTAAGTCAATGAGCGGGAGATGCCCTCAACGAGAACCTGGTTAAC 4368  
Db 4321 GGGCTGTCGGCTTTAAGTCAATGAGCGGGAGATGCCCTCAACGAGAACCTGGTTAAC 4380  
Qy 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTGGGGTCTGTGTCGCGAGCGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTGGGGTCTGTGTCGCGAGCGATA 4440  
Qy 4429 CTGCTGCGACGTGGGCCAGGGAGGGGGCTGTGCACTGGATGAACCGCTGTATAGCG 4488  
Db 4441 CTGCTGCGACGTGGGCCAGGGAGGGGGCTGTGCACTGGATGAACCGCTGTATAGCG 4500  
Qy 4489 TTGCGTTCCGGGGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCGAGCTGCA 4548  
Db 4501 TTGCGTTCCGGGGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCGAGCTGCA 4560  
Qy 4549 GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTGTCTGAAGGCTTCAC 4608  
Db 4561 GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTGTCTGAAGGCTTCAC 4620  
Qy 4609 CAGTGATCAACAGAGACTGCTCCACGCCATGCTCCGGTCTGFGTAAAGATGTTGG 4668  
Db 4621 CAGTGATCAACAGAGACTGCTCCACGCCATGCTCCGGTCTGFGTAAAGATGTTGG 4680  
Qy 4669 GATTGATATGACGCTGTGATGATTTCAAGCTGGCTCCAGTCCAGTCCCTCCCG 4728  
Db 4681 GATTGATATGACGCTGTGATGATTTCAAGCTGGCTCCAGTCCAGTCCCTCCCG 4740  
Qy 4729 CGATTCCCGGAGTCCCTCTCTCTCATGTCACAGTGGGTACAGGAGTCTGGCGGGC 4788  
Db 4741 CGATTCCCGGAGTCCCTCTCTCTCATGTCACAGTGGGTACAGGAGTCTGGCGGGC 4800  
Qy 4789 GACGGCATCATGCAAAACCACTCCCATGTCGAGCAAGATGTCAGGACATGTCGAGAAAC 4848  
Db 4801 GACGGCATCATGCAAAACCACTCCCATGTCGAGGCAAGATGTCAGGACATGTCGAGAAAC 4860  
Qy 4849 GGTTCATGAGATCGTGGGCTAGGACTGTAGTAAACAGTGGCATGGAACTTCC 4908  
Db 4861 GGTTCATGAGATCGTGGGCTAGGACTGTAGTAAACAGTGGCATGGAACTTCC 4920  
Qy 4909 ATTAAACGGGTACACCAAGCGCCCTGACGCCCTCCCGCGCCAAATATTCTAGGGG 4968  
Db 4921 ATTAAACGGGTACACCAAGCGCCCTGACGCCCTCCCGCGCCAAATATTCTAGGGG 4980  
Qy 4969 CTGTGCGGGTGTCTGCTGAGAGTACGTGAGAGTTACGCGGTGGGGATTTCCATC 5028  
Db 4981 CTGTGCGGGTGTCTGCTGAGAGTACGTGAGAGTTACGCGGTGGGGATTTCCATC 5040  
Qy 5029 GTGACGGGATGACCACTGACAACTGAAGTGGCCCGTGTGAGTTCGGGCCCGGAATTC 5088  
Db 5041 GTGACGGGATGACCACTGACAACTGAAGTGGCCCGTGTGAGTTCGGGCCCGGAATTC 5100  
Qy 5089 TTCAAGAGTGGATGGGGTGGCGTTGCCAGGTACGCTCCAGCGTGCAGAAACCCCTCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGCGTTGCCAGGTACGCTCCAGCGTGCAGAAACCCCTCTA 5160  
Qy 5149 CGGGAGAGTCACTTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGTCCCA 5208  
Db 5161 CGGGAGAGTCACTTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGTCCCA 5220

Qy 5209 TCGAGCCCAACCGACGTAGCGTCTCACTTCATGCTCACCGACCCCTCCACATT 5268  
Db 5221 TCGAGCCCAACCGACGTAGCGTCTCACTTCATGCTCACCGACCCCTCCACATT 5280  
Qy 5269 AGCGGGAGACGGCTAGGCTAGGCTGGCCAGGGATCTCCGCCCTCTCTGGCCAGCTCA 5328  
Db 5281 AGCGGGAGACGGCTAGGCTAGGCTGGCCAGGGATCTCTCCGCCCTCTGGCCAGCTCA 5340  
Qy 5329 TACGCTAGCAGCTGTCTGCGCTTCTCTTGAAGGCAACATGCACTACCGTCTAGCTCC 5388  
Db 5341 TACGCTAGCAGCTGTCTGCGCTTCTCTTGAAGGCAACATGCACTACCGTCTAGCTCC 5400  
Qy 5389 CGGAGCGCTGACTCATCGAGGCCACCTCTCTGTCGCGCAGAGATGGCGGGAACATC 5448  
Db 5401 CGGAGCGCTGACTCATCGAGGCCACCTCTCTGTCGCGCAGAGATGGCGGGAACATC 5460  
Qy 5449 ACCCGCTGAGTCAAGAAATTAAGGTAGTAAATTTTGGACTCTTTGAGCGCTCCAAAGCG 5508  
Db 5461 ACCCGCTGAGTCAAGAAATTAAGGTAGTAAATTTTGGACTCTTTGAGCGCTCCAAAGCG 5520  
Qy 5509 GAGGAGGATGAGGGAAGTATCCGTTCCGGCGGAGATCCTGCGGAGGTCAGGAAATTC 5568  
Db 5521 GAGGAGGATGAGGGAAGTATCCGTTCCGGCGGAGATCCTGCGGAGGTCAGGAAATTC 5580  
Qy 5569 CCTCGAGCGATGCCATATGGGCACGCCCGGATTACAAACCCTCACTGTTAGAGTCTCTGG 5628  
Db 5581 CCTCGAGCGATGCCATATGGGCACGCCCGGATTACAAACCCTCACTGTTAGAGTCTCTGG 5640  
Qy 5629 AAGGACCCGAGTACGTCCTCCAGTGTACAGGGTGTCCATTTGCGCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCCGAGTACGTCCTCCAGTGTGTACAGGGTGTCCATTTGCGCGCTGCCAAGGCC 5700  
Qy 5689 CCTCGATACCACTCCACGGAGGAAGAGACGGTTGTCCTGTCTCAGAAATCTACCGTCTCT 5748  
Db 5701 CCTCGATACCACTTCACGGAGGAAGAGACGGTTGTCCTGTCTCAGAAATCTACCGTCTCT 5760  
Qy 5749 TCTGCTTCCGCGAGTCCGCCAAGAGCTTCGCGAGCTCCGAATCTGTCGGCCGCTCGAC 5808  
Db 5761 TCTGCTTCCGCGAGTCCGCCAAGAGCTTCGCGAGCTTCGGAATCTGTCGGCCGCTCGAC 5820  
Qy 5809 AGCGGACCGGCAACCGCTCTCTGTACAGCCCTCCGACAGCGGACCGGAGATCCGAC 5868  
Db 5821 AGCGGACCGGCAACCGCTCTCTGTACAGCCCTCCGACAGCGGACCGGAGATCCGAC 5880  
Qy 5869 GTTGAGTCTACTCTCTCATGCCCCCTTGAAGGGGAGCGCGGGATCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTACTCTCTCATGCCCCCTTGAAGGGGAGCGCGGGATCCGATCTCAGC 5940  
Qy 5929 GACGGTCTTTGTTCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGCTCTCGATG 5988  
Db 5941 GACGGTCTTTGTTCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGCTCTCGATG 6000  
Qy 5989 TCTACATATGACAGCGCCCTGTATCAGCCATGCGCTGCGGAGAAACCAAGCTGCC 6048  
Db 6001 TCTACATATGACAGCGCCCTGTATCAGCCATGCGCTGCGGAGAAACCAAGCTGCC 6060  
Qy 6049 ATCAATGCACTGAGCAACTCTTTGTCCTCCATCACCACAACTTGCTCTATGCTCAACATCT 6108  
Db 6061 ATCAATGCACTGAGCAACTCTTTGTCCTCCATCACCACAACTTGCTCTATGCTCAACATCT 6120  
Qy 6109 CGCAGGCAAGCTGCGGACAGAAAGGTCACTTTTGAAGAGTCACTGAGGCTCTCTGACGAC 6168  
Db 6121 CGCAGGCAAGCTGCGGACAGAAAGGTCACTTTTGAAGAGTCACTGAGGCTCTCTGACGAC 6180  
Qy 6169 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAGGCTAACTT 6228  
Db 6181 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAGGCTAACTT 6240  
Qy 6229 CTATCCGTGAGGAAGCTGTAGCTGACGCCCACTTCGCGCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGAGGAAGCTGTAGCTGACGCCCACTTCGCGCAGATCTAAATTTGGC 6300



6289 TATGGGCAAGACGCTCGGAAACCTATCCAGCAAGCCGCTTAACACATCCCGTCG 6348  
6301 TATGGGCAAGACGCTCGGAAACCTATCCAGCAAGCCGCTTAACACATCCCGTCG 6360  
6349 TGAAGGACTTGTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAT 6408  
6361 TGAAGGACTTGTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAT 6420  
6409 GAGGTTTCTGCGTCCAAACAGAGAGGGGGGGGCGGAGCGAGTGCCTTATCGTATTC 6468  
6421 GAGGTTTCTGCGTCCAAACAGAGAGGGGGGGGCGGAGCGAGTGCCTTATCGTATTC 6480  
6469 CAGATTTGGGGGTTGTTGTCGCGAATAATGSCCTTTACGATGTTGCTCCACCTC 6528  
6481 CAGATTTGGGGGTTGTTGTCGCGAATAATGSCCTTTACGATGTTGCTCCACCTC 6540  
6529 CCTCAGGCGGTGATGGCTCTTATACGGAATCCAAATPACTCTCCTGACAGCGGTGAG 6588  
6541 CCTCAGGCGGTGATGGCTCTTATACGGAATCCAAATPACTCTCCTGACAGCGGTGAG 6600  
6589 TTCTGTTGATGCTGGAAGGAGGAATGCTTATGGCTTCGATATGACACCGC 6648  
6601 TTCTGTTGATGCTGGAAGGAGGAATGCTTATGGCTTCGATATGACACCGC 6660  
6649 TGTGTTGACTCAACGCTCAGTGAATGACATCGTGTGAGGAGTCAATCTACCAATG 6708  
6661 TGTGTTGACTCAACGCTCAGTGAATGACATCGTGTGAGGAGTCAATCTACCAATG 6720  
6709 TGTGACTTGGCCCCCGAAGCGACAGAGCCATTAAGTTCGCTCAGAGCGGCTTTATC 6768  
6721 TGTGACTTGGCCCCCGAAGCGACAGAGCCATTAAGTTCGCTCAGAGCGGCTTTATC 6780  
6769 GGGGGCCCCCTGACTTAATCTAAAGGCGAGAACTGCGGCTATCGCGTGCAGCGAGC 6828  
6781 GGGGGCCCCCTGACTTAATCTAAAGGCGAGAACTGCGGCTATCGCGTGCAGCGAGC 6840  
6829 GGTGTAAGTACGACACGCTCGGCTTAATPACCTCAGATGTTACTTGAAGGCGCTGCGGCC 6888  
6841 GGTGTAAGTACGACACGCTCGGCTTAATPACCTCAGATGTTACTTGAAGGCGCTGCGGCC 6900  
6889 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGCTT 6948  
6901 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGCTT 6960  
6949 ATCTGTGAAGCGGGGACCCAGAGAGAGAGCGAGCCCTACGGGCTTTCACGAGGCT 7008  
6961 ATCTGTGAAGCGGGGACCCAGAGAGAGAGCGAGCCCTACGGGCTTTCACGAGGCT 7020  
7009 ATGACTAGATCTCTGCCCCCTCGGGGACCCGCAAAACAGAAATACGACTTGGAGTTG 7068  
7021 ATGACTAGATCTCTGCCCCCTCGGGGACCCGCAAAACAGAAATACGACTTGGAGTTG 7080  
7069 ATAACATCATGCTCTTCAATGTTGCTAGTGCAGCAGCATGCTGTCGCAAAAGGTTGAC 7128  
7081 ATAACATCATGCTCTTCAATGTTGCTAGTGCAGCAGCATGCTGTCGCAAAAGGTTGAC 7140  
7129 TATCTACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7188  
7141 TATCTACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7200  
7189 ACTCCAGTCAATCTCTGGCTAGGCAACATCATATGATGTCGCCACCTTGTGGGCAAG 7248  
7201 ACTCCAGTCAATCTCTGGCTAGGCAACATCATATGATGTCGCCACCTTGTGGGCAAG 7260  
7249 ATGATCTGATGATCAATTTCTTCCATCTTCTAGCTCAGGAACCACTTGAAAAAGCC 7308  
7261 ATGATCTGATGATCAATTTCTTCTTCTTCTAGCTCAGGAACCACTTGAAAAAGCC 7320  
7309 CTGATTTGTCAGATCTACGGGCTGTTACTCCATGAGGACCTTACCTACCTCAGATC 7368  
7321 CTGATTTGTCAGATCTACGGGCTGTTACTCCATGAGGACCTTACCTACCTCAGATC 7380  
7369 ATTCAAGACTCCATGSCCTTAGGGCAATTTTCACTCCATAGTTACTCTCCAGGTGATC 7428

7381 ATTCAAGACTCCATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCAGGTGATC 7440  
7429 AATAGGTTGGCTTTCATGCTTACGCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACAT 7488  
7441 AATAGGTTGGCTTTCATGCTTACGCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACAT 7500  
7489 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGGAGGCTGCCACTTGTGGC 7548  
7501 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGGAGGCTGCCACTTGTGGC 7560  
7549 AAGTACTCTTCAACTTGGGCGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7608  
7561 AAGTACTCTTCAACTTGGGCGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7620  
7609 TCCAGTTGGAATTTATCCAGCTGGTTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7668  
7621 TCCAGTTGGAATTTATCCAGCTGGTTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7680  
7669 AGCTGTCTGTCGCGGACCCCGCTGGTTCACTGTGTGCTACTCCTACTTCTGTAGGG 7728  
7681 AGCTGTCTGTCGCGGACCCCGCTGGTTCACTGTGTGCTACTCCTACTTCTGTAGGG 7740  
7729 GTAGGCACTATCTACTTCCCAACCGATGAACGGGACCTTAAACACTCCAGGCGCAATAGG 7788  
7741 GTAGGCACTATCTACTTCCCAACCGATGAACGGGAGCTTAAACACTCCAGGCGCAATAGG 7800  
7789 CCAATCTGTTTTTTCCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTT 7848  
7801 CCAATCTGTTTTTTCCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTT 7860  
7849 TTTTCTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTAGC 7908  
7861 TTTTCTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTAGC 7920  
7909 CCTAGTCACGGCTAGCTGTGAAAGGTCGGTGAGCCGCTTGACTGCAGAGAGTGTGTATAC 7968  
7921 CCTAGTCACGGCTAGCTGTGAAAGGTCGGTGAGCCGCTTGACTGCAGAGAGTGTGTATAC 7980  
7969 TGGCCTCTCTGCAGATCAAGT 7989  
7981 TGGCCTCTCTGCAGATCAAGT 8001

RESULT 4  
US-09-539-601-16  
; Sequence 16, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/NS3-3'/9-13F  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
; OTHER INFORMATION: fusion protein  
; FEATURE:  
; NAME/KEY: RBS



; LOCATION: (1202)...(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)...(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural protein NS3-5B;
; OTHER INFORMATION: carries cell culture-adaptive mutations from clone
; OTHER INFORMATION: 9-13F
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (7771)...(8001)
US-09-539-601-16

Query Match 99.5%; Score 7946.2; DB 4; Length 8001;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 7976; Conservative 0; Mismatches 13; Indels 12; Gaps 1;

QY 1 GCCAGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
DB 1 GCCAGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60

QY 61 TCTTCACGAGAAAGCGTCTAGCCATGCGGTAGTATGATGCTGCTGAGGCTCCAGGAC 120
DB 61 TCTTCACGAGAAAGCGTCTAGCCATGCGGTAGTATGATGCTGCTGAGGCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATGCCAG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATGCCAG 180

QY 181 GACGACCGGGTCTCTTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGGGTGCCCC 240
DB 181 GACGACCGGGTCTCTTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGGGTGCCCC 240

QY 241 GCGAGCTCTGACCGAGTAGTGTGGTTCGCGAAAGCGTCTGCTGCTGATGCTGCTGATG 300
DB 241 GCGAGCTCTGACCGAGTAGTGTGGTTCGCGAAAGCGTCTGCTGCTGATGCTGCTGATG 300

QY 301 GTGCTTGGCGAGTCCCGGAGTCTGCTGAGCCGTCACATGAGCAGCAATCCGTAAC 360
DB 301 GTGCTTGGCGAGTCCCGGAGTCTGCTGAGCCGTCACATGAGCAGCAATCCGTAAC 360

QY 361 CTCGAAGAAAAACCAAA-----GGGCGCGCCATGATGAACAAGATGGATTGC 408
DB 361 CTCGAAGAAAAACCAAAACGTAACCAACGAGCGCGCCATGATGAACAAGATGGATTGC 420

QY 409 AGCGAGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGATGCGGCAACAGAG 468
DB 409 AGCGAGTTCTCGGCGCTTGGGTGAGAGGCTATTCGGCTATGATGCGGCAACAGAG 480

QY 469 CAATCGGCTGCTGATGCGCGCTGCTCCGCTGTCAGCGAGGGGCGCGGTTCTTT 528
DB 469 CAATCGGCTGCTGATGCGCGCTGCTCCGCTGTCAGCGAGGGGCGCGGTTCTTT 540

QY 529 TTGTCAAGACCGACTGTCGGTCCCTGAATCAACTGACGACGAGGCGCGGCTAT 588
DB 529 TTGTCAAGACCGACTGTCGGTCCCTGAATCAACTGACGACGAGGCGCGGCTAT 600

QY 589 CTTGCTGCGACGAGCGGCTCTTGGCGAGCTGCTGACAGTTGCTACTGAAGCGG 648
DB 589 CTTGCTGCGACGAGCGGCTCTTGGCGAGCTGCTGACAGTTGCTACTGAAGCGG 660

QY 649 GAAGGACTGGTCTGCTATTGGCGAGTCCCGGGGAGGATCTCTGCTATCTCACTTG 708
DB 649 GAAGGACTGGTCTGCTATTGGCGAGTCCCGGGGAGGATCTCTGCTATCTCACTTG 720

QY 709 CTCCTGCGAGAAAGTATCCATCATGGCTGATGCAATGGCGGCTGATACGCTGATC 768
DB 709 CTCCTGCGAGAAAGTATCCATCATGGCTGATGCAATGGCGGCTGATACGCTGATC 780

QY 769 CGGCTACCTGCCCATTCGACCCAGCCAGCAACATCGATCAGCGGACGCTACTCGGA 828
DB 769 CGGCTACCTGCCCATTCGACCCAGCCAGCAACATCGATCAGCGGACGCTACTCGGA 840

QY 829 TGGAGAGCCGGTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAG 888
DB 841 TGGAGAGCCGGTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAG 900

QY 889 CCGAACTGTTCCCGAGGCTCAAGGGCGCATCCCGAGCGGAGGATCTCTGCTGACCC 948
DB 901 CCGAACTGTTCCCGAGGCTCAAGGGCGCATCCCGAGCGGAGGATCTCTGCTGACCC 960

QY 949 ATGGCGATGCTGCTTGGCGGAATATCATGCTGGAAAAATGGCGCTTTTCTGATTCATCG 1008
DB 961 ATGGCGATGCTGCTTGGCGGAATATCATGCTGGAAAAATGGCGCTTTTCTGATTCATCG 1020

QY 1009 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCCGATGATA 1068
DB 1021 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCCGATGATA 1080

QY 1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTTCTCTGCTTTTACGATATCGCG 1128
DB 1081 TTGCTGAAGAGCTTGGCGGGAATGGGCTTCTCTGCTTTTACGATATCGCG 1140

QY 1129 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTTTAA 1188
DB 1141 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTTTAA 1200

QY 1189 CAGACCAACAAAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCT 1248
DB 1201 CAGACCAACAAAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCT 1260

QY 1249 AACGTTACTGCGGAGCGCTTGAATAAGCGCGGTGCTGCTTGTCTATATGTTT 1308
DB 1261 AACGTTACTGCGGAGCGCTTGAATAAGCGCGGTGCTGCTTGTCTATATGTTT 1320

QY 1309 TCACCATATTCGCGCTTTTGGCAATGTAGGCGCGGAAACCTGGCGCTCTCTCTTG 1368
DB 1321 TCACCATATTCGCGCTTTTGGCAATGTAGGCGCGGAAACCTGGCGCTCTCTCTTG 1380

QY 1369 AGGAGCTTCTTGGGCTTCTTCCCTCTCGCAAGGAAATCAAGGCTCTGTTGAATGTC 1428
DB 1381 AGGAGCTTCTTGGGCTTCTTCCCTCTCGCAAGGAAATCAAGGCTCTGTTGAATGTC 1440

QY 1429 GTGAAGGAGCAGTCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1488
DB 1441 GTGAAGGAGCAGTCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1500

QY 1489 TGCAGGCGCGGAAACCCCGCCTCGGACAGGTCCTCGGCGCAAAAGCCAGCTGTA 1548
DB 1501 TGCAGGCGCGGAAACCCCGCCTCGGACAGGTCCTCGGCGCAAAAGCCAGCTGTA 1560

QY 1549 TAAGATACACCTGCAAGGCGGCAACCCCGCCTCGGACAGGTCCTCGGCGCAAAAGCCAGCTGTA 1608
DB 1561 TAAGATACACCTGCAAGGCGGCAACCCCGCCTCGGACAGGTCCTCGGCGCAAAAGCCAGCTGTA 1620

QY 1609 GAAAGATCAATGCTCTCTCAAGCGTATTCACAAAGGCTGAAAGGATCCGAGAG 1668
DB 1621 GAAAGATCAATGCTCTCTCAAGCGTATTCACAAAGGCTGAAAGGATCCGAGAG 1680

QY 1669 GTACCCCATGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTACATGCTTTAG 1728
DB 1681 GTACCCCATGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTACATGCTTTAG 1740

QY 1729 TCGAGTTTAAAAAAGCTCTAGGCGCCCGCAACACGCGGAGCTGCTTTTCTTTGAAAA 1788
DB 1741 TCGAGTTTAAAAAAGCTCTAGGCGCCCGCAACACGCGGAGCTGCTTTTCTTTGAAAA 1800

QY 1789 CACGATATACCATGGCGCTTATACGGCTTCTCCCAACAGACGCGAGGCTCTTGCG 1848
DB 1801 CACGATATACCATGGCGCTTATACGGCTTCTCCCAACAGACGCGAGGCTCTTGCG 1860

QY 1849 TGCATCATCCTAGCTCAGAGCGGACAGAAACAGGTCGAGGGGAGGTCCAAAGTG 1908
DB 1861 TGCATCATCCTAGCTCAGAGCGGACAGAAACAGGTCGAGGGGAGGTCCAAAGTG 1920

QY 1909 GTCTCCACCGCAACACAATCTTTCTGCGCAGCTCGCTCAATGCGGTCTGTGGACTGTC 1968  
Db 1921 GTCTCCACCGCAACACAATCTTTCTGCGCAGCTCGCTCAATGCGGTCTGTGGACTGTC 1980  
QY 1969 TATCATGCTGCGGCTCAAGACACCTTTGCGGCGCCAAAGGCGCCCAATCACCCAAATGTAC 2028  
Db 1981 TATCATGCTGCGGCTCAAGACACCTTTGCGGCGCCAAAGGCGCCCAATCACCCAAATGTAC 2040  
QY 2029 ACCAATGTGACAGGAGCTCTGCTGCGGTGGGAAGGCGCCCGGCGCGGTTCCTTGACA 2088  
Db 2041 ACCAATGTGACAGGAGCTCTGCTGCGGTGGGAAGGCGCCCGGCGCGGTTCCTTGACA 2100  
QY 2089 CCATGCACTGCGGAGCTCGGACCTTTACTTTGTCTCAGAGGCATGCCGATCATTCGG 2148  
Db 2101 CCATGCACTGCGGAGCTCGGACCTTTACTTTGTCTCAGAGGCATGCCGATCATTCGG 2160  
QY 2149 GTGCGCGGCGGCGGCGAGCAGCAGGGGAGCCCTACTCTCCCCAGGCGCGCTCTCCTACTTG 2208  
Db 2161 GTGCGCGGCGGCGGCGAGCAGCAGGGGAGCCCTACTCTCCCCAGGCGCGCTCTCCTACTTG 2220  
QY 2209 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTCGGCGGCAAGCTGTGGGCATCTTCGG 2268  
Db 2221 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTCGGCGGCAAGCTGTGGGCATCTTCGG 2280  
QY 2269 GCTGCCGTGTGCAACCGAGGGTTTGGAGAGCGGTGGACTTTGTACCGCTCGACTCTATG 2328  
Db 2281 GCTGCCGTGTGCAACCGAGGGTTTGGAGAGCGGTGGACTTTGTACCGCTCGACTCTATG 2340  
QY 2329 GAAACCACTATGCGGTCCCGGCTTTCACGAGCAACTCGTCCCCCTCGGCGCGGTACCGCAG 2388  
Db 2341 GAAACCACTATGCGGTCCCGGCTTTCACGAGCAACTCGTCCCCCTCGGCGCGGTACCGCAG 2400  
QY 2389 ACATTCAGGTGGCCATCTACAGCCCTCTACTGGTAGCGGCAAGAGACTAAGGTGCG 2448  
Db 2401 ACATTCAGGTGGCCATCTACAGCCCTCTACTGGTAGCGGCAAGAGACTAAGGTGCG 2460  
QY 2449 GCTGCTATGCAAGGCTAAGGTGCTGCTGTAACCGCTCGCTGCGCGCCACC 2508  
Db 2461 GCTGCTATGCAAGGCTAAGGTGCTGCTGTAACCGCTCGCTGCGCGCCACC 2520  
QY 2509 CTAGGTTTCGGGCGTATATGCTTAAGGCACATGGTATCGACCCCTAACACAGAACCGGG 2568  
Db 2521 CTAGGTTTCGGGCGTATATGCTTAAGGCACATGGTATCGACCCCTAACACAGAACCGGG 2580  
QY 2569 GTAAGGACATACACCGGTGCCCCCATCAGCTACTCCACTATGCGAAGTTCCTGCG 2628  
Db 2581 GTAAGGACATACACCGGTGCCCCCATCAGCTACTCCACTATGCGAAGTTCCTGCG 2640  
QY 2629 GACGTTGTTGCTCTGCGGCGCCTATGACATCATATATATGTGATGAGTGCCACTCACT 2688  
Db 2641 GACGTTGTTGCTCTGCGGCGCCTATGACATCATATATATGTGATGAGTGCCACTCACT 2700  
QY 2689 GACTCGACCACTATCTTGGGCATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2748  
Db 2701 GACTCGACCACTATCTTGGGCATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2760  
QY 2749 CGACTCGTCTGCTCGCCACCGCTACGCCCTCGGGATCGGTACCGGTGCGACATPCAAAC 2808  
Db 2761 CGACTCGTCTGCTCGCCACCGCTACGCCCTCGGGATCGGTACCGGTGCGACATPCAAAC 2820  
QY 2809 ATCGAGAGTGGTCTGTCTGACGACCTGGAGAAATCCCTTTTATGGCAAGCCATCCC 2868  
Db 2821 ATCGAGAGTGGTCTGTCTGACGACCTGGAGAAATCCCTTTTATGGCAAGCCATCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAAAGATGTGAT 2928  
Db 2881 ATCGAGACCATCAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAAAGATGTGAT 2940  
QY 2929 GAGTCCCGGAGAGCTGTGCGGCTCGGACTCAATGCTGTAGCATATTAACGGGCGCTT 2988  
Db 2941 GAGTCCCGGAGAGCTATCGGCGCTCGGACTCAATGCTGTAGCATATTAACGGGCGCTT 3000  
QY 2989 GATGTATCCGTATACCACTAGCGGAGAGCTCATTTGCTGTAGCAACGGAGGCTCTAATG 3048

Db 3001 GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTGTCTGTAGCAACGGAGCTCTAATG 3060  
QY 3049 ACGGCTTTTACCGGCGATTTTCGACTCAGTGTATCAGTGTGCAATATGTGTCAACAGACA 3108  
Db 3061 ACGGCTTTTACCGGCGATTTTCGACTCAGTGTATCAGTGTGCAATATGTGTCAACAGACA 3120  
QY 3109 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3168  
Db 3121 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3180  
QY 3169 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3228  
Db 3181 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3240  
QY 3229 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3288  
Db 3241 GTGCACTTACGCTTGGACCCGACCTTCACTTGGAGCGAGCCGTTGCCACAGAGCGG 3300  
QY 3289 GACCGGCTGTGCTTGGTAGAGCTCAAGCTTCAAGCTTCCCTCGGTTCTGTGCGAGTGTAT 3348  
Db 3301 GACCGGCTGTGCTTGGTAGAGCTCAAGCTTCAAGCTTCCCTCGGTTCTGTGCGAGTGTAT 3360  
QY 3349 TACCTAACAACACAGAGGTTGCCGCTCTGCCAGAGCCATCTGGAGTTCCTGGAGAGCGTC 3408  
Db 3361 TACCTAACAACACAGAGGTTGCCGCTCTGCCAGAGCCATCTGGAGTTCCTGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTTCAACCATAGAGCGCCATTTCTTGTCCCAGACTAAGCAGGAGGAGAG 3468  
Db 3421 TTTACAGGCTTCAACCATAGAGCGCCATTTCTTGTCCCAGACTAAGCAGGAGGAGAG 3480  
QY 3469 AACTTCCCTTACCTGTGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTTCCACT 3528  
Db 3481 AACTTCCCTTACCTGTGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTTCCACT 3540  
QY 3529 CCATCGTGGAGCAAAATGTGGAAGTGTCTCATAGGCTAAGCTAAGCTACGCTGACGGGCCA 3588  
Db 3541 CCATCGTGGAGCAAAATGTGGAAGTGTCTCATAGGCTAAGCTAAGCTGACGGGCCA 3600  
QY 3589 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACACCCATA 3648  
Db 3601 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACACCCATA 3660  
QY 3649 ACCAAATACATATGGCATGTGCTGCTGAGCTTGGAGTGTGCTACGAGCACTGGGTG 3708  
Db 3661 ACCAAATACATATGGCATGTGCTGCTGAGCTGCTGAGGCTGCTACGAGCACTGGGTG 3720  
QY 3709 CTGTTAGCGGAGTCTCTAGCAGCTCTGGCGCGTATTTGCTGACACAGGAGCGGTGCTC 3768  
Db 3721 CTGTTAGCGGAGTCTCTAGCAGCTCTGGCGCGTATTTGCTGACACAGGAGCGGTGCTC 3780  
QY 3769 ATTGTGGGAGGATCATTTTGTCCGAAAGCCCGCCCATCATTTCCCGACAGGAGTCTT 3828  
Db 3781 ATTGTGGGAGGATCATTTTGTCCGAAAGCCCGCCCATCATTTCCCGACAGGAGTCTT 3840  
QY 3829 TACCGGAGTTCATGAGATGGAAGTGGCTCTCACCTCCCTTATCATGCAACAGGGA 3888  
Db 3841 TACCGGAGTTCATGAGATGGAAGTGGCTCTCACCTCCCTTATCATGCAACAGGGA 3900  
QY 3889 ATGCACTGCCGAACAAATTTCAACAGAGGCAATCGGGTTGCTGCAACAGCCCAAG 3948  
Db 3901 ATGCACTGCCGAACAAATTTCAACAGAGGCAATCGGGTTGCTGCAACAGCCCAAG 3960  
QY 3949 CAAGCGAGGCTCTGCTCTCCGTTGGAAATCCAAAGTGGCGGACCCCTCGAAGCTTCTG 4008  
Db 3961 CAAGCGAGGCTCTGCTCTCCGTTGGAAATCCAAAGTGGCGGACCCCTCGAAGCTTCTG 4020  
QY 4009 GCGAAGCATATGTGAATTTTCATCAGCGGATACAATTTTACAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGAATTTTCATCAGCGGATACAATTTTACAGGCTTGTCCACTCTG 4080  
QY 4069 CTTGGCAACCCCGGATAGCATCACTGATGGCATTCACAGCTCTATCATCAGCGGCTC 4128

Db	4081	CCTGGCAACCCCGGATAGCATCATCTGATGGCATTACAGCCCTCTATCACCAGCCGCTC	4140	QY	5209	TGCGAGCCGGAACCGSACGTAGCTGCTCACTTCATGCTCACGACCCCTCCCAATT	5268
QY	4129	ACCACCCACATACCTCTCTGTTTAACTCTTGGGGGATGGTGGCCGCCCAACTTGCT	4188	Db	5221	TCGGGGCCCCGAACCGGACGTAGCTGCTCACTTCATGCTCACGACCCCTCCCAATT	5280
Db	4141	ACCACCCACATACCTCTCTGTTTAACTCTTGGGGGATGGTGGCCGCCCAACTTGCT	4200	QY	5269	ACGGCGGAGACCGGCTAAAGCTAGGCTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA	5328
QY	4189	CCTCCAGCGCTGCTCTCTCTTCTGTPAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC	4248	Db	5281	ACGGCGGAGACCGGCTAAAGCTAGGCTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA	5340
Db	4201	CCTCCAGCGCTGCTCTCTCTTCTGTPAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC	4260	QY	5329	TCAGCTAGCAGCTGCTGGCCCTTCTTTGAAGGCAACATGCACTACCCGTATGACTCC	5388
QY	4249	ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGCAGGTATATGGAGCGGGTGGCAGGC	4308	Db	5341	TCAGCTAGCAGCTGCTGGCCCTTCTTTGAAGGCAACATGCACTACCCGTATGACTCC	5400
Db	4261	ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGCAGGTATATGGAGCGGGTGGCAGGC	4320	QY	5389	CGGAGCGTGACTCATCGAGGCCAACCTCTGTGGCGCAGAGATGGCGGGAACATC	5448
QY	4309	GGCTCTGTCGCTTTAAGTCAATGAGCGCGAGATGCCCTCCACCGAGACCTGTTAAC	4368	Db	5401	CGGAGCGTGACTCATCGAGGCCAACCTCTGTGGCGCAGAGATGGCGGGAACATC	5460
Db	4321	GGCTCTGTCGCTTTAAGTCAATGAGCGCGAGATGCCCTCCACCGAGACCTGTTAAC	4380	QY	5449	ACCCGCTGGAGTCAAGAAATAGGTAGTAAATTTTGGACTTTTGGAGCGCTCCAAAGCG	5508
QY	4369	CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCGAGGATA	4428	Db	5461	ACCCGCTGGAGTCAAGAAATAGGTAGTAAATTTTGGACTTTTGGAGCGCTCCAAAGCG	5520
Db	4381	CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCGAGGATA	4440	QY	5509	GAGGAGATGAGAGGAGTATCCGTTCCGGCGGAGATCCTCGGAGGTCCAGGAATTC	5568
QY	4429	CTGCTCGGACATGCTGGGCCAGGGAGGGGCTGTGCATGTGATGAACCGGCTGATAGCG	4488	Db	5521	GAGGAGATGAGAGGAGTATCCGTTCCGGCGGAGATCCTCGGAGGTCCAGGAATTC	5580
Db	4441	CTGCTCGGACATGCTGGGCCAGGGAGGGGCTGTGCATGTGATGAACCGGCTGATAGCG	4500	QY	5569	CCTCGAGCGATGCCATATGGGCACGCCCGGATTAACCCCTCCACTGTTAGAGTCTCTGG	5628
QY	4489	TTGCTTCCGGGGTAAACAGCTCTCCCGCAGCACTATGTGCCTGAGAGCAGCTGCA	4548	Db	5581	CCTCGAGCGATGCCATATGGGCACGCCCGGATTAACCCCTCCACTGTTAGAGTCTCTGG	5640
Db	4501	TTGCTTCCGGGGTAAACAGCTCTCCCGCAGCACTATGTGCCTGAGAGCAGCTGCA	4560	QY	5629	AAGGACCCGAGTACGTCTCTCCAGTGTACAGGGTTCATTTGCCGCTGCCAAGGCC	5688
QY	4549	GCACTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGTCTGTAAGAGCTTCAC	4608	Db	5641	AAGGACCCGAGTACGTCTCTCCAGTGTACAGGGTTCATTTGCCGCTGCCAAGGCC	5700
Db	4561	GCACTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGTCTGTAAGAGCTTCAC	4620	QY	5689	CCTCGATACCACTCTCCAGGAGGAGGACGGTGTCTGTGTCAGAACTACCCGTGTCT	5748
QY	4609	CAGTGGATCAACGAGACTGCTCCAGCCATGCTCCGGTCTGTTGCTAAGAGATGTTGG	4668	Db	5701	CCTCGATACCACTCTCCAGGAGGAGGACGGTGTCTGTGTCAGAACTACCCGTGTCT	5760
Db	4621	CAGTGGATCAACGAGACTGCTCCAGCCATGCTCCGGTCTGTTGCTAAGAGATGTTGG	4680	QY	5749	CTGCTCTGGCGAGCTCGCCCAAGAGCTTTCGGCAGCTCCGAAATCGTCGGCGCTCGAC	5808
QY	4669	GATTGATATGACGCTGTGATGATTTCAAGCTGCTCCAGTCCAGTCTCCGCG	4728	Db	5761	CTGCTCTGGCGAGCTCGCCCAAGAGCTTTCGGCAGCTCCGAAATCGTCGGCGCTCGAC	5820
Db	4681	GATTGATATGACGCTGTGATGATTTCAAGCTGCTCCAGTCCAGTCTCCGCG	4740	QY	5809	AGCGGACCGCAACCGCTCTCTGACAGCCCTCCGACGCGGACGCGGAGATCCGAC	5868
QY	4729	CGATTCCGGAGTCCCTCTCTCTCATGTCAAGCTGGGTACAAGGAGTCTTGGCGGGC	4788	Db	5821	AGCGGACCGCAACCGCTCTCTGACAGCCCTCCGACGCGGACGCGGAGATCCGAC	5880
Db	4741	CGATTCCGGAGTCCCTCTCTCTCATGTCAAGCTGGGTACAAGGAGTCTTGGCGGGC	4800	QY	5869	GTTGAGTCTGCTCTCTCATGCCCCCTTGGAGGGAGGACGGGGATCCCGATCTCAGC	5928
QY	4789	GACGGATCATGCAACACCTCCCATGTGGAGCACAGATCAACGGACATGTGAAAAAC	4848	Db	5881	GTTGAGTCTGCTCTCTCATGCCCCCTTGGAGGGAGGACGGGGATCCCGATCTCAGC	5940
Db	4801	GACGGATCATGCAACACCTCCCATGTGGAGCACAGATCAACGGACATGTGAAAAAC	4860	QY	5929	GACGGTCTTGGTCTACCGTAAGCAGGAGGCTTAGTGAGACGTCTGCTGCTCGATG	5988
QY	4849	GGTTCCATGAGATCGTGGGCTAGACCTGTAGTAAACAGTGGCATGGACATTCGCC	4908	Db	5941	GACGGTCTTGGTCTACCGTAAGCAGGAGGCTTAGTGAGACGTCTGCTGCTCGATG	6000
Db	4861	GGTTCCATGAGATCGTGGGCTAGACCTGTAGTAAACAGTGGCATGGACATTCGCC	4920	QY	5989	TCCTACATGACAGCGGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC	6048
QY	4909	ATTAAACCGGTACACACCGGCCCTCGACGCCCTCCCGCGCCAAATATTCTAGGGG	4968	Db	6001	TCCTACATGACAGCGGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC	6060
Db	4921	ATTAAACCGGTACACACCGGCCCTCGACGCCCTCCCGCGCCAAATATTCTAGGGG	4980	QY	6049	ATCAATGCACTGAGCACTCTTGTCTCGTCCAGTCCAGTCTTGTCTATGCTCAACATCT	6108
QY	4969	CTGTGGGGTGTGCTGTAGAGATACGTGGAGTTACGGGGTGGGGATTTCCACTAC	5028	Db	6061	GTCNAATGCACTGAGCACTCTTGTCTCGTCCAGTCCAGTCTTGTCTATGCTCAACATCT	6120
Db	4981	CTGTGGGGTGTGCTGTAGAGATACGTGGAGTTACGGGGTGGGGATTTCCACTAC	5040	QY	6109	CGCAGCGAAGCTCGCGGAGAGAGGTACCTTTTGACAGACTGCGAGTCTCTGGACGAC	6168
QY	5029	GTACCGGGATGACCACTGACACGTAAAGTCCCGCTGTCAGTTCCGGCCCCCGAATTC	5088	Db	6121	CGCAGCGAAGCTCGCGGAGAGAGGTACCTTTTGACAGACTGCGAGTCTCTGGACGAC	6180
Db	5041	GTACCGGGATGACCACTGACACGTAAAGTCCCGCTGTCAGTTCCGGCCCCCGAATTC	5100	QY	6169	CACTACCGGAGCGTCTCAAGGAGTGAAGCGGAGGCTCCACAGTTAAGGCTAACTT	6228
QY	5089	TTCAACAGAGTGGATGGGGTCCGCTGCAACAGTACGCTCCAGCGTGCACACCTCCTA	5148	Db	6181	CACTACCGGAGCGTCTCAAGGAGTGAAGCGGAGGCGTCCACAGTTAAGGCTAACTT	6240
Db	5101	TTCAACAGAGTGGATGGGGTCCGCTGCAACAGTACGCTCCAGCGTGCACACCTCCTA	5160	QY	6229	CTATCCGTGGAGGAGCTGTAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC	6288
QY	5149	CGGAGAGGATCACTTCTGCTGGGCTCAATCAATACCTGTTGGGTACAGTCCCA	5208	Db	6241	CTATCCGTGGAGGAGCTGTAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC	6300
Db	5161	CGGAGAGGATCACTTCTGCTGGGCTCAATCAATACCTGTTGGGTACAGTCCCA	5220				



```

; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural proteins N83 - 5B
; OTHER INFORMATION: of cell culture adapted clone no. 19
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (7771)..(8001)
; US-09-539-601-28

Query Match          99.4%; Score 7939.8; DB 4; Length 8001;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 7972; Conservative 0; Mismatches 17; Indels 12; Gaps 1;

QY 1 GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
DB 1 GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60

QY 61 TCTTCAAGCAAGAGGTCTAGCCATGCGGTAGTATGATGATGATGATGATGATGATGATGAT 120
DB 61 TCTTCAAGCAAGAGGTCTAGCCATGCGGTAGTATGATGATGATGATGATGATGATGATGAT 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180

QY 181 GAGACACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240
DB 181 GAGACACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240

QY 241 GCGAGACTGACCGGAGTAGTGTGGTTCGCGGAAGGCTTGTGGTACTGCTGATAGG 300
DB 241 GCGAGACTGACCGGAGTAGTGTGGTTCGCGGAAGGCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGGAGTCCCGGAGGCTCTGCTGAGCCGTGACCATGAGCAGCAATCTTAAC 360
DB 301 GTGCTTGGAGTCCCGGAGGCTCTGCTGAGCCGTGACCATGAGCAGCAATCTTAAC 360

QY 361 CTCAAGAAAAACCCAAA-----GGCGCGCCATGATTGAACAAGATGATTGC 408
DB 361 CTCAAGAAAAACCCAAA-----GGCGCGCCATGATTGAACAAGATGATTGC 420

QY 409 ACGCAGGTTCTCGCGCGGCTGGTGGTGGAGAGCTATTGGCTATGACTGGGCAACACAG 468
DB 409 ACGCAGGTTCTCGCGCGGCTGGTGGTGGAGAGCTATTGGCTATGACTGGGCAACACAG 480

QY 469 CAATCGGCTGCTCTGATGCGCGGCTGTTCCGGCTGTGACGCGAGGGCGCGGTTCTTT 528
DB 469 CAATCGGCTGCTCTGATGCGCGGCTGTTCCGGCTGTGACGCGAGGGCGCGGTTCTTT 540

QY 529 TTGTCAAGACCGACTGTCCGGTCCCTGGAATGAATGAACTGACGAGCAGGCGCGGCTAT 588
DB 529 TTGTCAAGACCGACTGTCCGGTCCCTGGAATGAATGAACTGACGAGCAGGCGCGGCTAT 600

QY 589 CGTGCTGCGCAGCAGCGGCTTCTTGGCAGCTGTGCTGACGCTGTCTGCTGCTGAGCGG 648
DB 589 CGTGCTGCGCAGCAGCGGCTTCTTGGCAGCTGTGCTGACGCTGTCTGCTGCTGAGCGG 660

QY 649 GAAGGACTGGTGTCTATTGGGCGAAGTGCCTGGGCGAGGATCTCCTGTGATCTCACTTTG 708
DB 649 GAAGGACTGGTGTCTATTGGGCGAAGTGCCTGGGCGAGGATCTCCTGTGATCTCACTTTG 720

QY 709 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGCGGGGTGATAGCTTGTATC 768
DB 709 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGCGGGGTGATAGCTTGTATC 780

QY 769 CGGCTACCTGCCCATTCGACACACCGAAGCAACATCGCATCGAGCGACCTACTCGGA 828
DB 769 CGGCTACCTGCCCATTCGACACACCGAAGCAACATCGCATCGAGCGACCTACTCGGA 840

```

```

QY 829 TCGAAGCCGGTCTTGTCTGATCAGGATGATCTGCAAGAGAGCATCAGGGGCTCGCCGAG 888
DB 841 TCGAAGCCGGTCTTGTCTGATCAGGATGATCTGCAAGAGAGCATCAGGGGCTCGCCGAG 900

QY 889 CCGAACTGTTCCCGAGGCTCAAGCGCGCATCCCGACGCGGAGGATCTCGTCTGACCC 948
DB 901 CCGAACTGTTCCCGAGGCTCAAGCGCGCATCCCGACGCGGAGGATCTCGTCTGACCC 960

QY 949 ATGGGATGCTGCTGCTGCGGAATATCATGCTGGAATAATGCGCGCTTTCTGATTCATCG 1008
DB 961 ATGGGATGCTGCTGCTGCGGAATATCATGCTGGAATAATGCGCGCTTTCTGATTCATCG 1020

QY 1009 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCGTGATA 1068
DB 1021 ACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCTGTGGCTACCGTGATA 1080

QY 1069 TTGCTGAAGAGCTTGGCGCGAATGGGTGACCGCTTCTCTGCTGCTTTACGATATGATA 1128
DB 1081 TTGCTGAAGAGCTTGGCGCGAATGGGTGACCGCTTCTCTGCTGCTTTACGATATGATA 1140

QY 1129 CTCGCGATTCGAGCGCATCGCTTCTATCGGCTTCTTACGAGTCTTCTGATTTAA 1188
DB 1141 CTCGCGATTCGAGCGCATCGCTTCTATCGGCTTCTTACGAGTCTTCTGATTTAA 1200

QY 1189 CAGACCAAAACCGTTTCCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCCCT 1248
DB 1201 CAGACCAAAACCGTTTCCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCCCT 1260

QY 1249 AACGTTACTGGCGAAGCGCTTGAATAAGCGCGGTGTGCTTGTCTATATGTTATTT 1308
DB 1261 AACGTTACTGGCGAAGCGCTTGAATAAGCGCGGTGTGCTTGTCTATATGTTATTT 1320

QY 1309 TCCACCATATGCGCTCTTTTGGCAATGTGAGGGCCGGAACCTGCGGCTCTGTTCTTG 1368
DB 1321 TCCACCATATGCGCTCTTTTGGCAATGTGAGGGCCGGAACCTGCGGCTCTGTTCTTG 1380

QY 1369 ACGAGCTTCTAGGGGTCTTTCCGCTCTCGCAAGGAATCAAGGTCTGTTGAATGTC 1428
DB 1381 ACGAGCTTCTAGGGGTCTTTCCGCTCTCGCAAGGAATCAAGGTCTGTTGAATGTC 1440

QY 1429 GTGAAGGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAACTGTGTAGCGACCTT 1488
DB 1441 GTGAAGGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAACTGTGTAGCGACCTT 1500

QY 1489 TCGCAGCAGCGAAGCCCGCCACCTCGCGACAGGTGCTCTGCGGCAAAAGCAAGTGA 1548
DB 1501 TCGCAGCAGCGAAGCCCGCCACCTCGCGACAGGTGCTCTGCGGCAAAAGCAAGTGA 1560

QY 1549 TAAGATACACCTGCAAGCGGCAAAACCCAGTGCACGTTGTGAGTTGATGTTG 1608
DB 1561 TAAGATACACCTGCAAGCGGCAAAACCCAGTGCACGTTGTGAGTTGATGTTG 1620

QY 1609 GAAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGGGCTCAAGATGCCAGAG 1668
DB 1621 GAAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGGGCTCAAGATGCCAGAG 1680

QY 1669 GTACCCCATCTGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTACATGCTTTAG 1728
DB 1681 GTACCCCATCTGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTACATGCTTTAG 1740

QY 1729 TCGAGTTAAAAAAGCTCTAGGCCCCCGCAACCAAGGGGAGTGTGTTTCTTTGAAAA 1788
DB 1741 TCGAGTTAAAAAAGCTCTAGGCCCCCGCAACCAAGGGGAGTGTGTTTCTTTGAAAA 1800

QY 1789 CACGATAATACATGGCGCTTATACGGCTTACTCCCAACAGACGCGAGGCTTCTGCG 1848
DB 1801 CACGATAATACATGGCGCTTATACGGCTTACTCCCAACAGACGCGAGGCTTCTGCG 1860

QY 1849 TGCATCATCTAGCTCACAGCGCGGACAGGAACAGGTCTGAGGGGAGTCTCAAGT 1908
DB 1861 TGCATCATCTAGCTCACAGCGCGGACAGGAACAGGTCTGAGGGGAGTCTCAAGT 1920

QY 1909 GTCTCCACCGCAACAAATCTTTCTGCGGACCTCGTCAATGGGCTGTGTTGGATGTC 1968

```









Db 6301 TATGGGCAAGAGCGTCCGGAACCTATCCAGCAAGCCGCTTACACATCCGCTCCGTC 6360  
Qy 6349 TGAAGGACTTGTGGAAGACACTGAGACACCAATGACACCAATCATGCGCAAAAAAT 6408  
Db 6361 TGAAGGACTTGTGGAAGACACTGAGACACCAATGACACCAATCATGCGCAAAAAAT 6420  
Qy 6409 GAGGTTTCTGCGTCCAAACAGAGAAAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6468  
Db 6421 GAGGTTTCTGCGTCCAAACAGAGAAAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6480  
Qy 6469 CCAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6528  
Db 6481 CCAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6540  
Qy 6529 CCTCAGGCGGTGATGGGCTCTTTCATAGGANTCCAAATCTCTCTGACAGCGGTCGAG 6588  
Db 6541 CCTCAGGCGGTGATGGGCTCTTTCATAGGANTCCAAATCTCTCTGACAGCGGTCGAG 6600  
Qy 6589 TTCTGTGTGAATGCTCGGAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGC 6648  
Db 6601 TTCTGTGTGAATGCTCGGAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGC 6660  
Qy 6649 TGTCTTCACTCAACCGGTCACTGAGAATGACATCCGTTGTAGGAGTCAATCTACCAATGT 6708  
Db 6661 TGTCTTCACTCAACCGGTCACTGAGAATGACATCCGTTGTAGGAGTCAATCTACCAATGT 6720  
Qy 6709 TGTGACTTGGCCCCCGAAGCCAGACAGGCCATAAGTGGTCTACAGAGCGGCTTTACATC 6768  
Db 6721 TGTGACTTGGCCCCCGAAGCCAGACAGGCCATAAGTGGTCTACAGAGCGGCTTTACATC 6780  
Qy 6769 GGGGGCCCCCTGACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6828  
Db 6781 GGGGGCCCCCTGACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6840  
Qy 6829 GGTGTACTGACGACGAGCTGGGTAAATACCTTCACATGTACTTTGAAGCGCGTGGCGCC 6888  
Db 6841 GGTGTACTGACGACGAGCTGGGTAAATACCTTCACATGTACTTTGAAGCGCGTGGCGCC 6900  
Qy 6889 TGTGAGCTCGAAGCTCCAGGACTGACGATGCTGATGCGGAGACGACCTTGTGCTT 6948  
Db 6901 TGTGAGCTCGAAGCTCCAGGACTGACGATGCTGATGCGGAGACGACCTTGTGCTT 6960  
Qy 6949 ATCTGTGAAGCGGGGAGCCAGAGAGAGAGCGAGCGGCTACGGGCTTACCGAGGCT 7008  
Db 6961 ATCTGTGAAGCGGGGAGCCAGAGAGAGAGAGCGAGCGGCTTACGGGCTTACCGAGGCT 7020  
Qy 7009 ATGACTAGATCTCTGCCCCCTGGGAGCCCGCCCAACAGAAATACGACTTGGAGTTG 7068  
Db 7021 ATGACTAGATCTCTGCCCCCTGGGAGCCCGCCCAACAGAAATACGACTTGGAGTTG 7080  
Qy 7069 ATAACTATGCTCTCAATGTGTAGTGTGCGGACGATGATGCGCAAAAGGCTGTAC 7128  
Db 7081 ATAACTATGCTCTCAATGTGTAGTGTGCGGACGATGATGCGCAAAAGGCTGTAC 7140  
Qy 7129 TATCTACCCGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACGCTAGACAC 7188  
Db 7141 TATCTACCCGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACGCTAGACAC 7200  
Qy 7189 ACTCCAGTCAANTCTGGCTAGGCAACATCATATGTATGCGCCACCTTGTGGGCAAGG 7248  
Db 7201 ACTCCAGTCAANTCTGGCTAGGCAACATCATATGTATGCGCCACCTTGTGGGCAAGG 7260  
Qy 7249 ATGATCTGATGATCAATTTCTTCTCCATGCTTCTAGCTCAGGAACAACTTGAAAGGCC 7308  
Db 7261 ATGATCTGATGATCAATTTCTTCTCCATGCTTCTAGCTCAGGAACAACTTGAAAGGCC 7320  
Qy 7309 CTAGATTGTGAGATCTAGGGGCTGTACTTCCATTGAGCCATTGACCTACCTCAGATC 7368  
Db 7321 CTAGATTGTGAGATCTAGGGGCTGTACTTCCATTGAGCCATTGACCTACCTCAGATC 7380  
Qy 7369 ATTCACGACTCCATGGCCCTTAGCGCATTTTCACTCCATGTTACTCTCCAGGTGATC 7428

Db 7381 ATTCACGACTCCATGGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGATC 7440  
Qy 7429 AATAGGTTGGCTTCAATGCCCTCAGGAACTTTGGGGTACCGCCCTTGCAGTCTGAGACAT 7488  
Db 7441 AATAGGTTGGCTTCAATGCCCTCAGGAACTTTGGGGTACCGCCCTTGCAGTCTGAGACAT 7500  
Qy 7489 CGGSCCAAGAGTGTCCGCGCTAGCTACTGTCCAGGGGGGGGGTGCACACTTGTGGC 7548  
Db 7501 CGGSCCAAGAGTGTCCGCGCTAGCTACTGTCCAGGGGGGGGGTGCACACTTGTGGC 7560  
Qy 7549 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7608  
Db 7561 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7620  
Qy 7609 TCCAGTTGGAATTTATCCAGCTGGTTCTGTGTGTACAGCGGGGAGACATATATCAC 7668  
Db 7621 TCCAGTTGGAATTTATCCAGCTGGTTCTGTGTGTGTACAGCGGGGAGACATATATCAC 7680  
Qy 7669 AGCTGTCTGTGCGCCGACCCGCTGTTCTATGTGGTGCCTACTCTCTCTTCTGTAGG 7728  
Db 7681 AGCTGTCTGTGCGCCGACCCGCTGTTCTATGTGGTGCCTACTCTCTCTTCTGTAGG 7740  
Qy 7729 GTAGGCATCTATCTACTCCCAACCCGATGAACGGGACCTAAACACATCCAGGCCAATAGG 7788  
Db 7741 GTAGGCATCTATCTACTCCCAACCCGATGAACGGGAGCTAAACACATCCAGGCCAATAGG 7800  
Qy 7789 CCAATCTGTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848  
Db 7801 CCAATCTGTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860  
Qy 7849 TTTTCTCTTTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7908  
Db 7861 TTTTCTCTTTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7920  
Qy 7909 CCTAGTCACGGCTAGCTGTGAAGGTCGAGCGGCTTGACCTGCAGAGAGTGTCTGATC 7968  
Db 7921 CCTAGTCACGGCTAGCTGTGAAGGTCGAGCGGCTTGACCTGCAGAGAGTGTCTGATC 7980  
Qy 7969 TGGCCTCTCTGCAGATCAAGT 7989  
Db 7981 TGGCCTCTCTGCAGATCAAGT 8001

RESULT 6  
US-09-539-601-4  
; Sequence 4, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; EARLIER FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 8637  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I377/NS2-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1181)  
; OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion  
; OTHER INFORMATION: protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1190)..(1800)



Db 1741 AACGTCTAGGCCCCCGAACCACGGGACGFGTTTTCCTTTTGAACACACGATAATACC 1800  
QY 1801 ATG----- 1803  
Db 1801 ATGGACGGGAGATGGCAGCATCGTGGAGGCGGGTTTTTCGTAGGTCTGATACTCTTG 1860  
QY 1804----- 1803  
Db 1861 ACCTTGTACCGCACTATAAGCTGTTCCTCGCTAGGCTCATATGGTGGTTACAATATTTT 1920  
QY 1804----- 1803  
Db 1921 ATCACAGGGCGAGGCACACTTGCAGTGTGGATCCCCCCCCCTCAAGCTTCGGGGGGG 1980  
QY 1804----- 1803  
Db 1981 CGGATGCCGTCACTCTCTCACTGCGCGATCCACCAGAGCTTAATCTTTTACCATCACC 2040  
QY 1804----- 1803  
Db 2041 AATACTTGTCTCGCATACTCGGTCCACTCATGTGTCTCAGGCTGGTATACCAAAAGTG 2100  
QY 1804----- 1803  
Db 2101 CCGTACTTCTGGCGCACACGGGCTCATTCGTGATGTCATGTCTGGTGGGAAGTTGCT 2160  
QY 1804----- 1803  
Db 2161 GGGGTCATTATGTCCAAATGGCTCTCATGAAGTTGGCCGCACTGACAGGTACGTACGTT 2220  
QY 1804----- 1803  
Db 2221 TATGACCATCTACCCCACTGCGGGAAGTGGGCCACCGGGCCCTACAGAGCTTGGGGTG 2280  
QY 1804----- 1803  
Db 2281 GCAGTTGAGCCCGTCTTCTCTGATATGAGACCAAGTTATACCTTGGGGGCGAGAC 2340  
QY 1804----- 1803  
Db 2341 ACCGGCGGTGGGGAATCATCTTGGGCTGCGCTCTCGGCCCGAGGGGAGGGAG 2400  
QY 1804----- 1804  
Db 2401 ATACATCTGGGACGGGACAGAGCTTGAAGGACGGGGTGGGACTCTCGGGCTTATT 2460  
QY 1813 ACGGCTTACTCCCAACAGACGGAGGCTACTTGGTGCATCATCACTAGCCTCACAGGC 1872  
Db 2461 ACGGCTTACTCCCAACAGACGGAGGCTACTTGGTGCATCATCACTAGCCTCACAGGC 2520  
QY 1873 CGGACAGGAAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCAACACAAATCTTTC 1932  
Db 2521 CGGACAGGAAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCAACACAAATCTTTC 2580  
QY 1933 CTGGCACTCTGCTCAATGCGGTGTGTGTGACTGTCTATCATGTGTCGGGCTCAAGAGCC 1992  
Db 2581 CTGGCACTCTGCTCAATGCGGTGTGTGTGACTGTCTATCATGTGTCGGGCTCAAGAGCC 2640  
QY 1993 CTTCGGGCGCCAAAGGCGCCAAATCACCCAAATGTACACCAATGTGACACAGGACCTCGTC 2052  
Db 2641 CTTCGGGCGCCAAAGGCGCCAAATCACCCAAATGTACACCAATGTGACACAGGACCTCGTC 2700  
QY 2053 GGTGCGCAAGGCGCCCGGGGCGGTCTCTTGTACACCATGTGACCTGTGGGCGAGCTGGAC 2112  
Db 2701 GGTGCGCAAGGCGCCCGGGGCGGTCTCTTGTACACCATGTGACACCATGTGACCTGTGGGCGAGCTGGAC 2760  
QY 2113 CTTTACTTGTTCAGAGGATGCGGTGTCTATTCGGTGTGCGGGGCGGCGAGCAGG 2172  
Db 2761 CTTTACTTGTTCAGAGGATGCGGTGTCTATTCGGTGTGCGGGGCGGCGAGCAGG 2820  
QY 2173 GGGAGCCTACTCTCCCCCAGGCGGCTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACTG 2232  
Db 2821 GGGAGCCTACTCTCCCCCAGGCGGCTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACTG 2880

QY 2233 CTCTGCCCCCTCGGGCACAGCTGTGGGCATCTTTTCGGGCTGCCGTGTGACCCCGAGGGTT 2292  
Db 2881 CTCTGCCCCCTCGGGCACAGCTGTGGGCATCTTTTCGGGCTGCCGTGTGACCCCGAGGGTT 2940  
QY 2293 GCGAAGCGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATGCGGTCCCGGTC 2352  
Db 2941 GCGAAGCGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATGCGGTCCCGGTC 3000  
QY 2353 TTACGAGCAACTCGTCCCTCCCGGCGGTACCGCAGACATTTCCAGGTGGCCCATCTACAC 2412  
Db 3001 TTACGAGCAACTCGTCCCTCCCGGCGGTACCGCAGACATTTCCAGGTGGCCCATCTACAC 3060  
QY 2413 GCCCCTACTGCTAGCGGCAAGACACTAAAGTGGCGGTGCGTATGAGGCCAAGGTTAT 2472  
Db 3061 GCCCCTACTGCTAGCGGCAAGACACTAAAGTGGCGGTGCGTATGAGGCCAAGGTTAT 3120  
QY 2473 AAGGTGCTTGTCTGAACCGTTCGTCGCGCCACCTAGGTTTCGGGGGCTATATGCT 2532  
Db 3121 AAGGTGCTTGTCTGAACCGTTCGTCGCGCCACCTAGGTTTCGGGGGCTATATGCT 3180  
QY 2533 AAGSCAATGCTATCGACCCCTAACATCAGAACCGGGGTAAAGGACCATCAACCGGTTGCC 2592  
Db 3181 AAGSCAATGCTATCGACCCCTAACATCAGAACCGGGGTAAAGGACCATCAACCGGTTGCC 3240  
QY 2593 CCATCAGTACTCCACCTATGGCAAGTTTCTTGGCGGACGGTGGTGTCTGGGGGCGCC 2652  
Db 3241 CCATCAGTACTCCACCTATGGCAAGTTTCTTGGCGGACGGTGGTGTCTGGGGGCGCC 3300  
QY 2653 TATGACATCATATATGTGATGATGCCACTCACTCACTGACCTCGACCATCTCTGGGATC 2712  
Db 3301 TATGACATCATATATGTGATGATGCCACTCACTCACTGACCTCGACCATCTCTGGGATC 3360  
QY 2713 GGCACATCTCGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTGCTCGCCACCGCT 2772  
Db 3361 GGCACATCTCGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTGCTCGCCACCGCT 3420  
QY 2773 ACGGCTCGGGATCGGTCAACCGTCCACATCCAAAATCGAGAGGTGGTCTGTCCAGC 2832  
Db 3421 ACGGCTCGGGATCGGTCAACCGTCCACATCCAAAATCGAGAGGTGGTCTGTCCAGC 3480  
QY 2833 ACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGG 2892  
Db 3481 ACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGG 3540  
QY 2893 CACTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCCCGGAAGCTGTCCGGC 2952  
Db 3541 CACTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCCCGGAAGCTGTCCGGC 3600  
QY 2953 CTGGACTCAATGCTGTAGCATATTTACCGGGGCTTGTATGATATCCGTCATACCAACTAGC 3012  
Db 3601 CTGGACTCAATGCTGTAGCATATTTACCGGGGCTTGTATGATATCCGTCATACCAACTAGC 3660  
QY 3013 GGAGAGCTATTCTGTAGCAACGGAGCTCTAATGACGGGCTTTACCGGCGATTTGAC 3072  
Db 3661 GGAGAGCTATTCTGTAGCAACGGAGCTCTAATGACGGGCTTTACCGGCGATTTGAC 3720  
QY 3073 TCAGTGATCGACTGCAATACATGTGTCAACCAAGAGTTCGACTTCAGCTTGGACCCGAGC 3132  
Db 3721 TCAGTGATCGACTGCAATACATGTGTCAACCAAGAGTTCGACTTCAGCTTGGACCCGAGC 3780  
QY 3133 TTCAACATTTAGACAGCAGCGGTGCCAAGAGCGCGGTGTCAAGCTTCGACGCGCGGAGGC 3192  
Db 3781 TTCAACATTTAGACAGCAGCGGTGCCAAGAGCGCGGTGTCAAGCTTCGACGCGCGGAGGC 3840  
QY 3193 AGGACTGGTAGGGGAGGATGGGCAATTTACAGGTTTGTACTCCAGAGAAAGCGGCTCG 3252  
Db 3841 AGGACTGGTAGGGGAGGATGGGCAATTTACAGGTTTGTACTCCAGAGAAAGCGGCTCG 3300  
QY 3253 GGCATGCTCGATTCTCGGTTCGTGCGAGGTCTATGACGCGGCTGTCTTGGTACGAG 3312  
Db 3901 GGCATGCTCGATTCTCGGTTCGTGCGAGGTCTATGACGCGGCTGTCTTGGTACGAG 3960

3313 CTACGCCGCCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACACACACAGGGTTGCC 3372  
3961 CTCACGCCGCCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACACACAGGGTTGCC 4020  
3373 GTCTCCAGGACCATCTGAGTTCTGGAGTCGTCTTTACAGGCTCAACCCACATAGAC 3432  
4021 GTCTCCAGGACCATCTGAGTTCTGGAGTCGTCTTTACAGGCTCAACCCACATAGAC 4080  
3433 GCCCATTTCTTGTCCAGACTAAGCAGCAGAGAGCAACTTCCCTCCTGTAGCATAC 3492  
4081 GCCCATTTCTTGTCCAGACTAAGCAGCAGAGAGCAACTTCCCTCCTGTAGCATAC 4140  
3493 CAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCTGCGGACCAAAATGTGGAAG 3552  
4141 CAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCTGCGGACCAAAATGTGGAAG 4200  
3553 TGCTCATACGCTAAAGCTACGCTGCAACGCGCCAAAGCCCTGCTGTATAGGCTGGGA 3612  
4201 TGCTCATACGCTAAAGCTACGCTGCAACGCGCCAAAGCCCTGCTGTATAGGCTGGGA 4260  
3613 GCGGTTCAAAACGAGTTACTACCAACACACCCCATTAACCAATACATCATGCGCATG 3672  
4261 GCGGTTCAAAACGAGTTACTACCAACACACCCCATTAACCAATACATCATGCGCATG 4320  
3673 TCGGCTGACCTGAGGTCGTACAGACACCTGGGTCGTGTTAGGCGGAGTCTTAGAGCT 3732  
4321 TCGGCTGACCTGAGGTCGTACAGACACCTGGGTCGTGTTAGGCGGAGTCTTAGAGCT 4380  
3733 CTGGCGCGGTATGCTGACAAACAGGACAGCGTGTCAATTTGGGCGAGGATCATCTTGTCC 3792  
4381 CTGGCGCGGTATGCTGACAAACAGGACAGCGTGTCAATTTGGGCGAGGATCATCTTGTCC 4440  
3793 GGAAGCGCGCATCATCTCCGACAGGAGTCTTTACCGGAGTTCGATGAGATGGA 3852  
4441 GGAAGCGCGCATCATCTCCGACAGGAGTCTTTACCGGAGTTCGATGAGATGGA 4500  
3853 GAGTGGCTCACACCTCCTTACATCGAACAGGAGTGCAGCTCGCCGAAACAAATCAAA 3912  
4501 GAGTGGCTCACACCTCCTTACATCGAACAGGAGTGCAGCTCGCCGAAACAAATCAAA 4560  
3913 CAGAAGGCAATCGGTTGTGCAAAACAGCACAAGCAAGCGGAGGCTGCTGCCGCTG 3972  
4561 CAGAAGGCAATCGGTTGTGCAAAACAGCACAAGCAAGCGGAGGCTGCTGCCGCTG 4620  
3973 GTGGATCAAGTGGCGGACCTCGAAGCTTCTGGCGAGCATATGTTGATTCATC 4032  
4621 GTGGATCAAGTGGCGGACCTCGAAGCTTCTGGCGAGCATATGTTGATTCATC 4680  
4033 AGCGGATCAATATTTAGCAGCTTGTCCACTCTGCTGCTGGCAACCCCGGATAGCATCA 4092  
4681 AGCGGATCAATATTTAGCAGCTTGTCCACTCTGCTGCTGGCAACCCCGGATAGCATCA 4740  
4093 CTGATGGCATTCACAGCTTATCAACAGCCGCTCAGCACCACCAACATACCTCTGTTT 4152  
4741 CTGATGGCATTCACAGCTTATCAACAGCCGCTCAGCACCACCAACATACCTCTGTTT 4800  
4153 AACATCTGGGGGATGGTGGCGCCCACTTGTCTCCAGCGCTGCTCTGCTTTC 4212  
4801 AACATCTGGGGGATGGTGGCGCCCACTTGTCTCCAGCGCTGCTCTGCTTTC 4860  
4213 GTAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTGTG 4272  
4861 GTAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTGTG 4920  
4273 GATATTTGGCAGTTATGAGCAGGGTGGCAGGCGCTGTTGGCTTTAGGTCATG 4332  
4921 GATATTTGGCAGTTATGAGCAGGGTGGCAGGCGCTGTTGGCTTTAGGTCATG 4980  
4333 AGCGGCGAGATGCCCTCCACCGAGGACCTGTTTAAACCTACTCCTGCTATCTCTCCCT 4392  
4981 AGCGGCGAGATGCCCTCCACCGAGGACCTGTTTAAACCTACTCCTGCTATCTCTCCCT 5040  
4393 GGCGCCCTAGTCTCGGGGTCTGTGCGCAGCATATGCTGCGGACGTTGGGCCACAGG 4452

5041 GGCGCCCTAGTCTCGGGTCTGTGCGCAGCATACTCTGCTCGGCACGTTGGGCCACAGG 5100  
4453 GAGGGGGTGTGCTAGTGAACCGGCTGATAGGTTCTGCTTCCGGGGTAAACACAGTC 4512  
5101 GAGGGGGTGTGCTAGTGAACCGGCTGATAGGTTCTGCTTCCGGGGTAAACACAGTC 5160  
4513 TCCGCCACGCACTATGTGCTGAGAGCGACGCTGACGACACGCTGTCACTCAGATCTCTCT 4572  
5161 TCCGCCACGCACTATGTGCTGAGAGCGACGCTGACGACACGCTGTCACTCAGATCTCTCT 5220  
4573 AGCTTTACATCACTCAGCTGCTGAAGGCTTCAACAGTGGATCAACGAGACTGCTCC 4632  
5221 AGCTTTACATCACTCAGCTGCTGAAGGCTTCAACAGTGGATCAACGAGACTGCTCC 5280  
4633 AGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGATGATGACAGTGTGACT 4692  
5281 AGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGATGATGACAGTGTGACT 5340  
4693 GATTTCAAGACCTGGCTCCAGTCCAGCTCTGCGCGGATTTCCCGGAGTCCCTTCTTC 4752  
5341 GATTTCAAGACCTGGCTCCAGTCCAGCTCTGCGCGGATTTCCCGGAGTCCCTTCTTC 5400  
4753 TCATGTCAACGTTGGGTAAAGGGAGTCTGGCGGGCGACGGCATCATGCAAAACCACTGC 4812  
5401 TCATGTCAACGTTGGGTAAAGGGAGTCTGGCGGGCGACGGCATCATGCAAAACCACTGC 5460  
4813 CCATGTGAGACACATCACCGACATGTGAACACGTTTCCATGAGGATCGTGGGCT 4872  
5461 CCATGTGAGACACATCACCGACATGTGAACACGTTTCCATGAGGATCGTGGGCT 5520  
4873 AGGACCTGTAGTAAACACGCTGGATGGAACATTTCCCATTAACCGGCTACACCAAGGCCCC 4932  
5521 AGGACCTGTAGTAAACACGCTGGATGGAACATTTCCCATTAACCGGCTACACCAAGGCCCC 5580  
4933 TGACGCGCTTCCCGGCGCAATTTATTTAGGGGCTGTGGCGGTGCTGCTGAGGAG 4992  
5581 TGACGCGCTTCCCGGCGCAATTTATTTAGGGGCTGTGGCGGTGCTGCTGAGGAG 5640  
4993 TAGTGGAGTTACCGGGTGGGATTTCCACTAGTCACGGCATGACCACTGACAAAC 5052  
5641 TAGTGGAGTTACCGGGTGGGATTTCCACTAGTCACGGCATGACCACTGACAAAC 5700  
5053 GTAAAGTGGCTGTGCTGAGTTCTCGGCCCCCGAATTTCTTACAGAAAGTGGATGGGTGCGG 5112  
5701 GTAAAGTGGCTGTGCTGAGTTCTCGGCCCCCGAATTTCTTACAGAAAGTGGATGGGTGCGG 5760  
5113 TTGCAAGATACGCTCAGCTGCAAAACCCCTCTCTAGGGAGGATGACATTTCTTCTGTC 5172  
5761 TTGCAAGATACGCTCAGCTGCAAAACCCCTCTCTAGGGAGGATGACATTTCTTCTGTC 5820  
5173 GGGCTCAATCAATACCTGTTGGGTCAACAGCTCCCATGCGAGCCGAAACGGACGTAGCA 5232  
5821 GGGCTCAATCAATACCTGTTGGGTCAACAGCTCCCATGCGAGCCGAAACGGACGTAGCA 5880  
5233 GTGCTCACTTCCATGCTCAGACCCCTCCCATTTACCGCGGAGACGGCTAAGCGTAGG 5292  
5881 GTGCTCACTTCCATGCTCAGACCCCTCCCATTTACCGCGGAGACGGCTAAGCGTAGG 5940  
5293 CTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCATAGCTAGCCAGCTGTCTGCGCT 5352  
5941 CTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCATAGCTAGCCAGCTGTCTGCGCT 6000  
5353 TCCCTTGAAGCAACATGCACTACCCCTCATGCTCCCGGAGCGCTGACCTCATCGAGGCC 5412  
6001 TCCCTTGAAGCAACATGCACTACCCCTCATGCTCCCGGAGCGCTGACCTCATCGAGGCC 6060  
5413 AACCTCTGTGGCGGAGATGCGCGGAAACATCAACCGCGTGGAGTGAAGATAAG 5472  
6061 AACCTCTGTGGCGGAGATGCGCGGAAACATCAACCGCGTGGAGTGAAGATAAG 6120  
5473 GTAGTAAATTTGACCTTTTCGAGCGCTCCAGCGGAGGATGAGAGGATATCC 5532

Db 6121 GTAGTAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGAGGAGGATGAGAGGAAAGTATCC 6180  
Qy 5533 GTTCCGCGGAGATCCTCGGAGGTCCAGGAAATTCCTTCAGCGATGCCCATATGGGCA 5592  
Db 6181 GTTCCGCGGAGATCCTCGGAGGTCCAGGAAATTCCTTCAGCGATGCCCATATGGGCA 6240  
Qy 5593 CGCCGGATTAACAACTCTCCACTGTTAGAGTCTTGGAAGGACCGGACTACGTCCCTCCA 5652  
Db 6241 CGCCGGATTAACAACTCTCCACTGTTAGAGTCTTGGAAGGACCGGACTACGTCCCTCCA 6300  
Qy 5653 GTGGTACACGGGTGTCATTCGCGCTGCGAAGGCCCTCCGATACCACTCCACGGAGG 5712  
Db 6301 GTGGTACACGGGTGTCATTCGCGCTGCGAAGGCCCTCCGATACCACTCCACGGAGG 6360  
Qy 5713 AAGAGGACGGTGTCTGTGTCAGAAATCTACCGTGTCTTCGCTTGGCGGAGCTCGGCACA 5772  
Db 6361 AAGAGGACGGTGTCTGTGTCAGAAATCTACCGTGTCTTCGCTTGGCGGAGCTCGGCACA 6420  
Qy 5773 AAGACCTTCGCGAGCTCCGAATCTCGGCCCTCGACAGCGCACCGCAACGGCTCTCCT 5832  
Db 6421 AAGACCTTCGCGAGCTCCGAATCTCGGCCCTCGACAGCGCACCGCAACGGCTCTCCT 6480  
Qy 5833 GACAGCCCTCCGACGACGCGGATCCGAGCTTGAGTCTGATCTCTCCATGCC 5892  
Db 6481 GACAGCCCTCCGACGACGCGGATCCGAGCTTGAGTCTGATCTCTCCATGCC 6540  
Qy 5893 CCCCTTGAGGGGAGCGGGGATCCGATCTCAGCGACGGTCTTGGTCTACCGTAAGC 5952  
Db 6541 CCCCTTGAGGGGAGCGGGGATCCGATCTCAGCGACGGTCTTGGTCTACCGTAAGC 6600  
Qy 5953 GAGGAGCTAGTGAGGACGTCGTCTGCTGCTCGATGTCCTACATGAGCAGGGCGCTG 6012  
Db 6601 GAGGAGCTAGTGAGGACGTCGTCTGCTGCTCGATGTCCTACATGAGCAGGGCGCTG 6660  
Qy 6013 ATACGCGCATGCGTGGGAGGAAACCAAGTGTCCCATCAATGCACTGAGCAACTCTTTG 6072  
Db 6661 ATACGCGCATGCGTGGGAGGAAACCAAGTGTCCCATCAATGCACTGAGCAACTCTTTG 6720  
Qy 6073 CTCGTCACACACTTGTGTATGCTACACATCTCGACGCGCAAGCTCGGCGAAG 6132  
Db 6721 CTCGTCACACACTTGTGTATGCTACACATCTCGACGCGCAAGCTCGGCGAAG 6780  
Qy 6133 AAGGTCACTTTGACAGACTGCAAGTCTCGACGACCACTACCGGACGTCCTCAAGGAG 6192  
Db 6781 AAGGTCACTTTGACAGACTGCAAGTCTCGACGACCACTACCGGACGTCCTCAAGGAG 6840  
Qy 6193 ATGAAGCGAAGCGTCCACAGTTAAGCTTAACTTCTATCCGTTGAGGAGCGCTGTAA 6252  
Db 6841 ATGAAGCGAAGCGTCCACAGTTAAGCTTAACTTCTATCCGTTGAGGAGCGCTGTAA 6900  
Qy 6253 CTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGCAAGGACGTCGCGAAC 6312  
Db 6901 CTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGCAAGGACGTCGCGAAC 6960  
Qy 6313 CTATCCAGCAAGCGTTAAACCATCCGCTCGGTGTGGAAGGACTTGTCTGGAAGACACT 6372  
Db 6961 CTATCCAGCAAGCGTTAAACCATCCGCTCGGTGTGGAAGGACTTGTCTGGAAGACACT 7020  
Qy 6373 GAGACCAATTAACACCACTATGGCAAAAATGAGGTTTCTGCGTCCAAACAGAG 6432  
Db 7021 GAGACCAATTAACACCACTATGGCAAAAATGAGGTTTCTGCGTCCAAACAGAG 7080  
Qy 6433 AAGGGGGCGCAGCGAGCTCGCTTATCGTATTCAGATTTGGGGGTTCGTGTGTC 6492  
Db 7081 AAGGGGGCGCAGCGAGCTCGCTTATCGTATTCAGATTTGGGGGTTCGTGTGTC 7140  
Qy 6493 GAGAAATGCCCCTTTACGATGTGCTCCACCTCCCTCAGCGCGTGATGGGCTCTTCA 6552  
Db 7141 GAGAAATGCCCCTTTACGATGTGCTCCACCTCCCTCAGCGCGTGATGGGCTCTTCA 7200  
Qy 6553 TAGCGATTCCAATCTCTCTGGACAGCGGTGAGTCTCTGTTGATGCTGGAAAGCG 6612  
Db 7201 TAGCGATTCCAATCTCTCTGGACAGCGGTGAGTCTCTGTTGATGCTGGAAAGCG 7260

Qy 6613 AAGAAATGCCCCTATGGGCTTCGATATGACACCCGCTGTTTGTGACTCAACGGTCACTGAG 6672  
Db 7261 AAGAAATGCCCCTATGGGCTTCGATATGACACCCGCTGTTTGTGACTCAACGGTCACTGAG 7320  
Qy 6673 AATGACATCCGTTGAGGAGTCAATCTACCAATGTTGACTTGGCCCCCGAAGCCAGA 6732  
Db 7321 AATGACATCCGTTGAGGAGTCAATCTACCAATGTTGACTTGGCCCCCGAAGCCAGA 7380  
Qy 6733 CAGGCCATAAGGTGCTTCACAGAGCGCTTTACATCGGGGGCCCCCTGACTAAATTTCTAAA 6792  
Db 7381 CAGGCCATAAGGTGCTTCACAGAGCGCTTTACATCGGGGGCCCCCTGACTAAATTTCTAAA 7440  
Qy 6793 GGGCAGAACTGCGGCTATCGCCGTCGCGCGAGCGGTGACTGACGACGAGCTGCGGT 6852  
Db 7441 GGGCAGAACTGCGGCTATCGCCGTCGCGCGAGCGGTGACTGACGACGAGCTGCGGT 7500  
Qy 6853 AATACCTCTCACATGTTACTTGAAGCGCTGCGGCTGCTGAGCTGCGAAGCTCCAGAC 6912  
Db 7501 AATACCTCTCACATGTTACTTGAAGCGCTGCGGCTGCTGAGCTGCGAAGCTCCAGAC 7560  
Qy 6913 TGCACGATGCTCGTATCGGAGACGACTTGTCTGTTATCTGTGAAAGCGGGACCCAA 6972  
Db 7561 TGCACGATGCTCGTATCGGAGACGACTTGTCTGTTATCTGTGAAAGCGGGACCCAA 7620  
Qy 6973 GAGGACGAGCGAGCTTACGGGCTTTCAGGAGGCTATGACTAGATATCTGCCCCCT 7032  
Db 7621 GAGGACGAGCGAGCTTACGGGCTTTCAGGAGGCTATGACTAGATATCTGCCCCCT 7680  
Qy 7033 GGGGACCCGCCCAACAGAAATCGACTTGGAGTTGATTAACATCATCTCTCCAATGTG 7092  
Db 7681 GGGGACCCGCCCAACAGAAATCGACTTGGAGTTGATTAACATCATCTCTCCAATGTG 7740  
Qy 7093 TCAGTCCGCGACGATGCTTGGCAAGGCTGACTATCTCACCCGCTGACCCACCCACC 7152  
Db 7741 TCAGTCCGCGACGATGCTTGGCAAGGCTGACTATCTCACCCGCTGACCCACCCACC 7800  
Qy 7153 CCCCTTGCGGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAATTCCTGGCTAGG 7212  
Db 7801 CCCCTTGCGGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAATTCCTGGCTAGG 7860  
Qy 7213 AACATCATCATGATGCGCCACCTTGTGGCAGGATGATCTCTGATGACTCATTTCTTC 7272  
Db 7861 AACATCATCATGATGCGCCACCTTGTGGCAGGATGATCTCTGATGACTCATTTCTTC 7920  
Qy 7273 TCCATCTCTTAGCTCAGGAACAACTTGAAGGCTTAGATTTGTAGATCTCCTGCTAGG 7332  
Db 7921 TCCATCTCTTAGCTCAGGAACAACTTGAAGGCTTAGATTTGTAGATCTCCTGCTAGG 7980  
Qy 7333 TGTATCTCATGAGGACCTGACCTACCTCAGATCAATCAAGACTCCATGGCTTAGC 7392  
Db 7981 TGTATCTCATGAGGACCTGACCTACCTCAGATCAATCAAGACTCCATGGCTTAGC 8040  
Qy 7393 GCATTTTCACTCCATAGTTACTCTCCAGGTGAGTCAATAGGTTGGCTTATGCTCAGG 7452  
Db 8041 GCATTTTCACTCCATAGTTACTCTCCAGGTGAGTCAATAGGTTGGCTTATGCTCAGG 8100  
Qy 7453 AAACTTTGGGTCACCGCTTTCGAGTCTGGAGACATCGGGCCAGAAAGTTCCTGGCTAGG 7512  
Db 8101 AAACTTTGGGTCACCGCTTTCGAGTCTGGAGACATCGGGCCAGAAAGTTCCTGGCTAGG 8160  
Qy 7513 CTAATGTCACGAGGGGAGGGCTGCACTTGTGGAAGTACCTCTTCACTGGGAGTA 7572  
Db 8161 CTAATGTCACGAGGGGAGGGCTGCACTTGTGGAAGTACCTCTTCACTGGGAGTA 8220  
Qy 7573 AGGACCAAGCTCAACTCTCAATCCGCTGCTCCAGTTGATTTATCCAGCTGG 7632  
Db 8221 AGGACCAAGCTCAACTCTCAATCCGCTGCTCCAGTTGATTTATCCAGCTGG 8280  
Qy 7633 TTGTTGCTGCTTACAGCGGGGAGACATATATCAAGGCTGCTCTGCTGCCGACCCCGC 7692  
Db 8281 TTGTTGCTGCTTACAGCGGGGAGACATATATCAAGGCTGCTCTGCTGCCGACCCCGC 8340



1009 ACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATA 1068  
1021 ACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATA 1080  
1069 TTGCTGAAGAGCTTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGATTCGCGG 1128  
1081 TTGCTGAAGAGCTTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGATTCGCGG 1140  
1129 CTCCCGATTGCGAGCGGATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1188  
1141 CTCCCGATTGCGAGCGGATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1200  
1189 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248  
1201 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1260  
1249 AAGCTTACTGGCGGAGCGCTTTGGAATAAGGCGCGGTGCGTTGTGCTATATGTTATTT 1308  
1261 AAGCTTACTGGCGGAGCGCTTTGGAATAAGGCGCGGTGCGTTGTGCTATATGTTATTT 1320  
1309 TCACCATATTTGCGCTTTTGTGCAATGTGAGGCGCGGAAACCTGCGCCCTGTTCTTTG 1368  
1321 TCACCATATTTGCGCTTTTGTGCAATGTGAGGCGCGGAAACCTGCGCCCTGTTCTTTG 1380  
1369 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTTGAATGTC 1428  
1381 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTTGAATGTC 1440  
1429 GTGAAGGAACAGTTCCTCTGGAAGTCTTTGAAGACAAACACCTCTGAGCGACCCCT 1488  
1441 GTGAAGGAACAGTTCCTCTGGAAGTCTTTGAAGACAAACACCTCTGAGCGACCCCT 1500  
1489 TGACGAGCGGGAACCCCTGAGTCTGCGGACAGGTGCTCTGCGGCCAAAAGCCACGTGTA 1548  
1501 TGACGAGCGGGAACCCCTGAGTCTGCGGACAGGTGCTCTGCGGCCAAAAGCCACGTGTA 1560  
1549 TAAGATACCTGCAAGGGGCGACACCCAGTGCACCGTCTGAGTCTGAGTGTGATGTTG 1608  
1561 TAAGATACCTGCAAGGGGCGACACCCAGTGCACCGTCTGAGTCTGAGTGTGATGTTG 1620  
1609 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1668  
1621 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1680  
1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTAATGTTTATG 1728  
1681 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTAATGTTTATG 1740  
1729 TCAGAGTTAAAAACGTTAGGCGCCCGGAAACACCGGAGCGTGTTCCTTTGAAAAA 1788  
1741 TCAGAGTTAAAAACGTTAGGCGCCCGGAAACACCGGAGCGTGTTCCTTTGAAAAA 1800  
1789 CAGATAATAACCATG----- 1803  
1801 CAGATAATAACCATGACCGGAGATGGAGCATCGTGGGAGCGCGGTTTTCGTAGGT 1860  
1804 ----- 1803  
1861 CTGATACTCTTGACCTTGTACCGCATATAAGTGTTCCTCGTAGGCTCATATGGTGG 1920  
1804 ----- 1803  
1921 TTACAATATTTTATCAGAGGCGGAGGACACATTGCAAGTGTGGATTCGCCCCCTCAAC 1980  
1804 ----- 1803  
1981 GTTCGGGGGGCGCGATGCGGTATCCTCCTCAGGTGCGCGATCCACCAGAGCTAATC 2040  
1804 ----- 1803  
2041 TTTTACCATCACCACAAATCTTGTCGCCATCTCGGTCCACTCATGTGTCCAGGCTGTT 2100

1804 ----- 1803  
2101 ATAACAAAGTGGCGTATCTTGTGCGGACACAGGGGCTCATCTGTGATGATGCTGGTG 2160  
1804 ----- 1803  
2161 CGGAAGTTTCTGGGGTCAATTATGTCCAATGGCTCTCATGAAGTTGGCCGCACTGACA 2220  
1804 ----- 1803  
2221 GSTACGTACGTTTATGACCATCTCACCCACTGCGGACTGGGCCACGCGGGCCTACGA 2280  
1804 ----- 1803  
2281 GACCTTGGCGTGCAGTTGAGCCCGTGTCTTCTCTGATATGGAGACCAAGTTTATCACC 2340  
1804 ----- 1803  
2341 TGGGGGACAGACACCGCGCGGTGTGGGGACATCATCTTGGGCCCTGCCCTCTCGCCCGC 2400  
1804 ----- 1803  
2401 AGGGGAGGAGATACATCTGGGACCGGACAGACAGCCCTTGAAGGCGAGGGTGGCGACTC 2460  
1804 ----- 1803  
2461 CTGCGCCCTATTACGGCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCTCATCACT 2520  
1861 AGCTCACAGGCGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
2521 AGCTCACAGGCGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA 2580  
1921 ACACAATCTTCTGCGGACCTCGCTCAATGGCGTGTGTTGAGTGTCTATCATGTGTC 1980  
2581 ACACAATCTTCTGCGGACCTCGCTCAATGGCGTGTGTTGAGTGTCTATCATGTGTC 2640  
1981 GGCTCAAGAGACCTTGGCGGCCCAAGGGGCCAATCACCCAAATGTACCAATGTGGAC 2040  
2641 GGCTCAAGAGACCTTGGCGGCCCAAGGGGCCAATCACCCAAATGTACCAATGTGGAC 2700  
2041 CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCTTGCACCATGCACTGTC 2100  
2701 CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCTTGCACCATGCACTGTC 2760  
2101 GGAGCTGCGACCTTTTACTTGTGTCAGAGGATGCGGATGCTATTCGGTSCCGCGCGG 2160  
2761 GGAGCTGCGACCTTTTACTTGTGTCAGAGGATGCGGATGCTATTCGGTSCCGCGCGG 2820  
2161 GGCGACAGAGGGGAGCGCTTCTCTCCCGAGCGCGTCTCTACTTGAAGGCTCTTTCG 2220  
2821 GGCGACAGAGGGGAGCGCTTCTCTCCCGAGCGCGTCTCTACTTGAAGGCTCTTTCG 2880  
2221 GGCGGTCACCTGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCGGTGTC 2280  
2881 GGCGGTCACCTGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCGGTGTC 2840  
2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTATCCCGTCGAGTCTATGGAACCACTATG 2340  
2941 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTATCCCGTCGAGTCTATGGAACCACTATG 3000  
2341 CGGTCCCGGCTTTCAGGACAACTCGTCCCTCGCGCGTACCGGACGATTCAGGTTG 2400  
3001 CGGTCCCGGCTTTCAGGACAACTCGTCCCTCGCGCGTACCGGACGATTCAGGTTG 3060  
2401 GCCCATCTACGCCCTTACTGTAGCGCAAGAGCACTAAAGTGCCTGCGGTGGTATGCA 2460  
3061 GCCCATCTACGCCCTTACTGTAGCGCAAGAGCACTAAAGTGCCTGCGGTGGTATGCA 3120  
2461 GCCCAAGGGTATAAGGTGCTTGTCTGAACCGGTCCGTCGCCGCCACCCCTAGGTTTCGGG 2520  
3121 GCCCAAGGGTATAAGGTGCTTGTCTGAACCGGTCCGTCGCCGCCACCCCTAGGTTTCGGG 3180  
2521 GCGTATATCTTAAGGCACATGTATCGACCCCTAAATCATCAGAACCGGGGTAAAGGACATC 2580





QY 4741 GTCCCCCTTCTTCTCATGTCAAAGTGGGTACAAAGGAGTCTGGGGGCGAGCGCATCATG 4800  
 Db 5401 GTCCCCCTTCTTCTCATGTCAAAGTGGGTACAAAGGAGTCTGGGGGCGAGCGCATCATG 5460  
 QY 4801 CAAACACCTGCCATGTGGAGACACAGATCACCGGACATGTGAAAAAGGTTCCATGAGG 4860  
 Db 5461 CAAACACCTGCCATGTGGAGACACAGATCACCGGACATGTGAAAAAGGTTCCATGAGG 5520  
 QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAAATTCGCCATTAACGCGTAC 4920  
 Db 5521 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAAATTCGCCATTAACGCGTAC 5580  
 QY 4921 ACCACGGCCCTCGAGCCCTCCCGGGCCCAATTAATCTAGGGCGCTGTGGCGGTG 4980  
 Db 5581 ACCACGGCCCTCGAGCCCTCCCGGGCCCAATTAATCTAGGGCGCTGTGGCGGTG 5640  
 QY 4981 GCTGCTGAGGAGTACGTGGAGGTATAGCGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
 Db 5641 GCTGCTGAGGAGTACGTGGAGGTATAGCGGGTGGGGATTTCCACTACGTGACGGGCATG 5700  
 QY 5041 ACCACTGACAACTAAAGTCCCGTGTCAAGTTCGGGCCCGCCGAAATTTCTCAAGAGTG 5100  
 Db 5701 ACCACTGACAACTAAAGTCCCGTGTCAAGTTCGGGCCCGCCGAAATTTCTCAAGAGTG 5760  
 QY 5101 GATGGGTGGGTGACAGGTACGTCCAGCGTGCAGACCCCTCTCAAGGAGGAGTTC 5160  
 Db 5761 GATGGGTGGGTGACAGGTACGTCCAGCGTGCAGACCCCTCTCAAGGAGGAGTTC 5820  
 QY 5161 ACATTCCTGTCGGGTCAATCAATACCTGGTGGGTACAGCTCCCATCGAGGCCGAA 5220  
 Db 5821 ACATTCCTGTCGGGTCAATCAATACCTGGTGGGTACAGCTCCCATCGAGGCCGAA 5880  
 QY 5221 CCGGACGTAGAGTGTCACTTCCATGCTCAACGACCCCTCCACATTAACGCGGAGAG 5280  
 Db 5881 CCGGACGTAGAGTGTCACTTCCATGCTCAACGACCCCTCCACATTAACGCGGAGAG 5940  
 QY 5281 GCTAAGCGTAGGCTGCGAGGATCTCCCGCTCTTGGCAGCTCATCAGTAGCCAG 5340  
 Db 5941 GCTAAGCGTAGGCTGCGAGGATCTCCCGCTCTTGGCAGCTCATCAGTAGCCAG 6000  
 QY 5341 CTGCTGCGCTTCCCTTGAAGGCAACATGACATACCTGCTCATGCTCCCGGAGCTGAC 5400  
 Db 6001 CTGCTGCGCTTCCCTTGAAGGCAACATGACATACCTGCTCATGCTCCCGGAGCTGAC 6060  
 QY 5401 CTATCGAGCCAACTCTCTGCGCGGAGAGATGGCGGGAACATCAACCCGCTGGAG 5460  
 Db 6061 CTATCGAGCCAACTCTCTGCGCGGAGAGATGGCGGGAACATCAACCCGCTGGAG 6120  
 QY 5461 TCAGAAATTAAGTAGTAATTTGGACTCTTCGAGCGCTTCCAGGGAGGAGTAG 5520  
 Db 6121 TCAGAAATTAAGTAGTAATTTGGACTCTTCGAGCGCTTCCAGGGAGGAGTAG 6180  
 QY 5521 AGGGAAGTATCCGCTCCGCGGAGATCTCGGAGGTCCAGGAAATTCCTCCAGCGATG 5580  
 Db 6181 AGGGAAGTATCCGCTCCGCGGAGATCTCGGAGGTCCAGGAAATTCCTCCAGCGATG 6240  
 QY 5581 CCCATATGGGACCGCCGGATTAACACCTCCACTGTAGAGTCTCTGGAAGGACCCGGAC 5640  
 Db 6241 CCCATATGGGACCGCCGGATTAACACCTCCACTGTAGAGTCTCTGGAAGGACCCGGAC 6300  
 QY 5641 TAGCTCCCTCCAGTGTACAGGCTGTCATTCGCGCTGCGAGGCCCTCCGATACCA 5700  
 Db 6301 TAGCTCCCTCCAGTGTACAGGCTGTCATTCGCGCTGCGAGGCCCTCCGATACCA 6360  
 QY 5701 CCTCCAGGAGGAGGAGCGGTGCTCTGCTAGAAATCTACCGTGTCTTCTGCTGGCG 5760  
 Db 6361 CCTCCAGGAGGAGGAGCGGTGCTCTGCTAGAAATCTACCGTGTCTTCTGCTGGCG 6420  
 QY 5761 GAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCGCGCTCGACAGCGGACGGA 5820  
 Db 6421 GAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCGCGCTCGACAGCGGACGGA 6480

QY 5821 AGCGCTCTCTGACAGACCCCTCCGAGCGGCGAGCGCGGATCCGAGCTTGAGTGTGAC 5880  
 Db 6481 AGCGCTCTCTGACAGACCCCTCCGAGCGGCGAGCGCGGATCCGAGCTTGAGTGTGAC 6540  
 QY 5881 TCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGCGGACGGGTCTTGG 5940  
 Db 6541 TCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGCGGACGGGTCTTGG 6600  
 QY 5941 TCTAACCGTAAAGGAGGAGCTAGTAGAGAGCTGTGTGTCTGATGTCTTCAATGAG 6000  
 Db 6601 TCTAACCGTAAAGGAGGAGCTAGTAGAGAGCTGTGTGTCTGATGTCTTCAATGAG 6660  
 QY 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGAAACAAAGCTGCCATCAATGACATG 6060  
 Db 6661 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGAAACAAAGCTGCCATCAATGACATG 6720  
 QY 6061 AGCAACTCTTTGCTCGCTCACCACAACTTGGTCTATGTCTACAACTCTCGACGCAAGC 6120  
 Db 6721 AGCAACTCTTTGCTCGCTCACCACAACTTGGTCTATGTCTACAACTCTCGACGCAAGC 6780  
 QY 6121 GTGGGCGAAGAGTCACTTTTGACAGACTGCGAGTCTGTGACAGCACTACCGGAC 6180  
 Db 6781 GTGGGCGAAGAGTCACTTTTGACAGACTGCGAGTCTGTGACAGCACTACCGGAC 6840  
 QY 6181 GTGCTCAAGGAGATGAAGCGGCTCCACAGTTAAGSCTAAACTTCTATCCGTGGAG 6240  
 Db 6841 GTGCTCAAGGAGATGAAGCGGCTCCACAGTTAAGSCTAAACTTCTATCCGTGGAG 6900  
 QY 6241 GAAGCTGTAAAGCTGACGCCCCCACAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6300  
 Db 6901 GAAGCTGTAAAGCTGACGCCCCCACAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6960  
 QY 6301 GAGCTCCGAAACTTATCCAGCAAGCGGTAAACCACTCCGCTCCGTGTGGAGAGCTTG 6360  
 Db 6961 GAGCTCCGAAACTTATCCAGCAAGCGGTAAACCACTCCGCTCCGTGTGGAGAGCTTG 7020  
 QY 6361 CTGGAGACACTGAGACCAATTTGACACCACTCATCGGCAAAATGAGTTTCTGC 6420  
 Db 7021 CTGGAGACACTGAGACCAATTTGACACCACTCATCGGCAAAATGAGTTTCTGC 7080  
 QY 6421 GTCCAAACGAGAGAGGGGCGCGAGCCAGCTCGGCTTATCGTATTTCCAGATTTGGGG 6480  
 Db 7081 GTCCAAACGAGAGAGGGGCGCGAGCCAGCTCGGCTTATCGTATTTCCAGATTTGGGG 7140  
 QY 6481 GTTCTGTGTGCGAGAAATATGCGCTTTACAGTGTGTCTCCACCTCCCTCAGGCGGTG 6540  
 Db 7141 GTTCTGTGTGCGAGAAATATGCGCTTTACAGTGTGTCTCCACCTCCCTCAGGCGGTG 7200  
 QY 6541 ATGGGCTCTTCAACGATTTCCAAATCTCTCTGGACAGGGGTGAGTTCCTGCTGAAT 6600  
 Db 7201 ATGGGCTCTTCAACGATTTCCAAATCTCTCTGGACAGGGGTGAGTTCCTGCTGAAT 7260  
 QY 6601 GCTGGAAGCGAAGAAATGCGCTTATGGGCTTCCGATATGACACCGGCTGTTTGAATCA 6660  
 Db 7261 GCTGGAAGCGAAGAAATGCGCTTATGGGCTTCCGATATGACACCGGCTGTTTGAATCA 7320  
 QY 6661 AGGTCACTGAGAAATGACATTCGCTGTGTGAGAGTCAATCTCAATGTGTGACTGGCC 6720  
 Db 7321 AGGTCACTGAGAAATGACATTCGCTGTGTGAGAGTCAATCTCAATGTGTGACTGGCC 7380  
 QY 6721 CCGGAGCCAGACAGCCATAGGTGCTCAGAGAGGCTTTACATCGGGGCCCCCTG 6780  
 Db 7381 CCGGAGCCAGACAGCCATAGGTGCTCAGAGAGGCTTTACATCGGGGCCCCCTG 7440  
 QY 6781 ACTAATTTTAAAGGCGAGAACTGCGCTATCGCGGTGCGCGGAGCGGTGTACTGACG 6840  
 Db 7441 ACTAATTTTAAAGGCGAGAACTGCGCTATCGCGGTGCGCGGAGCGGTGTACTGACG 7500  
 QY 6841 ACCAGTCCGGTAATACCTCTCACTGTATTCTTAAGGCGGCTGCGGCTGTGAGCTGGG 6900  
 Db 7501 ACCAGTCCGGTAATACCTCTCACTGTATTCTTAAGGCGGCTGCGGCTGTGAGCTGGG 7560  
 QY 6901 AAGCTCCAGGACTGCGAGGATGCTGTATCGGAGAGCACTTGTCTGTGAAAGC 6960

Db 7561 AAGCTCCAGGACTGCAACGATGCTCGTATGCGAGAGACCTTGTGCTTATCTGTGAAGC 7620  
QY 6961 GCGGGAGCCCAAGAGGAGGAGCGAGCTTA CGGGCTTTCACGGAGGCTATGACTAGATAC 7020  
Db 7621 GCGGGAGCCCAAGAGGAGGAGCGAGCTTA CGGGCTTTCACGGAGGCTATGACTAGATAC 7680  
QY 7021 TCTGCCCCCTGGGGAGCCCGCCAAACAGAAATACGACTTGGAGTGTGATACATCATGC 7080  
Db 7681 TCTGCCCCCTGGGGAGCCCGCCAAACAGAAATACGACTTGGAGTGTGATACATCATGC 7740  
QY 7081 TCTCCAATGTCTAGTCGCGACGATGATCTGGCAAAAGGCTGTACTTCTCAACCGT 7140  
Db 7741 TCTCCAATGTCTAGTCGCGACGATGATCTGGCAAAAGGCTGTACTTCTCAACCGT 7800  
QY 7141 GACCCACACACCCCTTGGCGGCTGGTGGGAGACGCTAGACACACTCCAGTCAAT 7200  
Db 7801 GACCCACACACCCCTTGGCGGCTGGTGGGAGACGCTAGACACACTCCAGTCAAT 7860  
QY 7201 TCTGGCTAGGCAACATCATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260  
Db 7861 TCTGGCTAGGCAACATCATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7920  
QY 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGA AAAAGCCCTAGATTGTGAG 7320  
Db 7921 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGA AAAAGCCCTAGATTGTGAG 7980  
QY 7321 ATCTACGGGGCTGTACTTCCATGAGGACCTTGTAGCTACCTACCTACATCATCAAGACTC 7380  
Db 7981 ATCTACGGGGCTGTACTTCCATGAGGACCTTGTAGCTACCTACCTACATCATCAAGACTC 8040  
QY 7381 CATGGCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT 7440  
Db 8041 CATGGCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT 8100  
QY 7441 TCATGCTCAGGAACCTTGGGCTACCGCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500  
Db 8101 TCATGCTCAGGAACCTTGGGCTACCGCTTGGAGTCTGGAGACATCGGGCCAGAAGT 8160  
QY 7501 GTCCGCGCTAGCTACTGTCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 8161 GTCCGCGCTAGCTACTGTCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 8220  
QY 7561 AACTGGGAGTAAGACCAAGCTCAAACTCACTCCAACTCCGCGTCCGCTCCAGTTGGAT 7620  
Db 8221 AACTGGGAGTAAGACCAAGCTCAAACTCACTCCAACTCCGCGTCCGCTCCAGTTGGAT 8280  
QY 7621 TTATCCAGCTGGTTCGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680  
Db 8281 TTATCCAGCTGGTTCGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 8340  
QY 7681 GCCGACCCCGCTGTTTCACTGTTGCTTACTCTTCTTCTGTTAGGGGTAGGATCTAT 7740  
Db 8341 GCCGACCCCGCTGTTTCACTGTTGCTTACTCTTCTTCTGTTAGGGGTAGGATCTAT 8400  
QY 7741 CTACTTCCCAACCGATGAACGGGACCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 7800  
Db 8401 CTACTTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 8460  
QY 7801 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db 8461 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8520  
QY 7861 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db 8521 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8580  
QY 7921 TAGCTGTGAAGGTCGGTGAGCGGCTTGA CTGAGAGAGTGTGATAGTGGCTCTCTGC 7980  
Db 8581 TAGCTGTGAAGGTCGGTGAGCGGCTTGA CTGAGAGAGTGTGATAGTGGCTCTCTGC 8640  
QY 7981 AGATCAAGT 7989  
|||||

Db 8641 AGATCAAGT 8649  
RESULT 8  
US-10-029-907-1  
; Sequence 1, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: HEPATITIS C VIRUS  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 8639  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1803) ... (8408)  
US-10-029-907-1  
Query Match 91.0%; Score 7272.2; DB 4; Length 8639;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 7975; Conservative 0; Mismatches 3; Indels 671; Gaps 3;  
QY 1 GCCAGCCCCGATTCGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60  
Db 2 GCCAGCCCCGATTCGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 61  
QY 61 TCTTCACGAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGTCAGGCTCCAGGAC 120  
Db 62 TCTTCACGAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGTCAGGCTCCAGGAC 121  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCGCGAG 180  
Db 122 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCGCGAG 181  
QY 181 GACGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTCGAGATTTGGGCGTCCGCC 240  
Db 182 GACGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTCGAGATTTGGGCGTCCGCC 241  
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCCCTTGTGTTACTGCTCTGATAGG 300  
Db 242 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCCCTTGTGTTACTGCTCTGATAGG 301  
QY 301 GTGCTTGGAGTGCCCGGGAGGTCTCGTAGCCGTGACCATGAGCAGCATCTTAAC 360  
Db 302 GTGCTTGGAGTGCCCGGGAGGTCTCGTAGCCGTGACCATGAGCAGCATCTTAAC 361  
QY 361 CTCAGAAAGAAACAAAGGGCGCGCATGATTGAACAAGATGGATTGCACGAGGTTCTC 420  
Db 362 CTCAGAAAGAAACAAAGGGCGCGCATGATTGAACAAGATGGATTGCACGAGGTTCTC 421  
QY 421 CGGCGCTTGGTGGAGAGGCTATTCGGCTATGACTGGGCAACAACAGACAATTCGGCTGT 480  
Db 422 CGGCGCTTGGTGGAGAGGCTATTCGGCTATGACTGGGCAACAACAGACAATTCGGCTGT 481  
QY 481 CTGATGCGCGTGTTCGGCTGTGACGCGAGGGGCGCGGCTTTTGTTCAGAGCG 540  
Db 482 CTGATGCGCGTGTTCGGCTGTGACGCGAGGGGCGCGGCTTTTGTTCAGAGCG 541  
QY 541 ACCTGTCGCGTCCCTGAATCAATGTCAGGACGAGGCGCGGCTATCGTGGCTGGCCA 600  
Db 542 ACCTGTCGCGTCCCTGAATCAATGTCAGGACGAGGCGCGGCTATCGTGGCTGGCCA 601  
QY 601 CGACGGGCTTCTTTCGCGAGCTGTGTCGAGCTTGTCACTGAAGGGGAGGAGCTGSC 660  
|||||

Db 602 CGAGGGCGCTTCTTGGCGAGCTGTGCTCGACGTCTCACTGAAGCGGAGGACTGGC 661  
QY 661 TGGTATTGGGCGAAGTGGCGGGGACAGATCTCTGTCACTCTCACTTGTCTCTTGGCGGAGA 720  
Db 662 TGGTATTGGGCGAAGTGGCGGGGACAGATCTCTGTCACTCTCACTTGTCTCTTGGCGGAGA 721  
QY 721 AAGTATCCATCATGGCTGATGCAATCGGCGGCTGCATAGCTTGATCGGCTTACCTGGC 780  
Db 722 AAGTATCCATCATGGCTGATGCAATCGGCGGCTGCATAGCTTGATCGGCTTACCTGGC 781  
QY 781 CATTCGACCACCAAGCGAATCGCATCGAGCGAGCAGTACTCGGATGGAAGCCGGTC 840  
Db 782 CATTCGACCACCAAGCGAATCGCATCGAGCGAGCAGTACTCGGATGGAAGCCGGTC 841  
QY 841 TTGTCGATCAGGATGATCTGGAGGAAGAGCATCAGGGGCTCGGCCAGCGGAATCTTTCG 900  
Db 842 TTGTCGATCAGGATGATCTGGAGGAAGAGCATCAGGGGCTCGGCCAGCGGAATCTTTCG 901  
QY 901 CCAGGCTCAAGGCGGCATGCCGACGCGAGGATCTCGTCTGACCCATGGCGATGCCCT 960  
Db 902 CCAGGCTCAAGGCGGCATGCCGACGCGAGGATCTCGTCTGACCCATGGCGATGCCCT 961  
QY 961 GCTTGCCTGAATATCATGCTGGAAATGGCGCTTTTCTGGAATCATCGACTGTGGCGGC 1020  
Db 962 GCTTGCCTGAATATCATGCTGGAAATGGCGCTTTTCTGGAATCATCGACTGTGGCGGC 1021  
QY 1021 TGGGTTGGCGGACCGCTATCAGGATAGCGTTGGCTTACCGTGTATTTCTGGAAGGC 1080  
Db 1022 TGGGTTGGCGGACCGCTATCAGGATAGCGTTGGCTTACCGTGTATTTCTGGAAGGC 1081  
QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGTGCTTTACGGTATCGCGCTTCCCGATTTCG 1140  
Db 1082 TTGGCGGCGAATGGGCTGACCGCTTCTGTGCTTTACGGTATCGCGCTTCCCGATTTCG 1141  
QY 1141 AGCGATCGCTTCTATCGCTTCTTGAAGGATCTTCTGAGTT-----TAAA 1188  
Db 1142 AGCGATCGCTTCTATCGCTTCTTGAAGGATCTTCTGAGTTTCGCGCCAGATGTTAA 1201  
QY 1189 CAGACCAACGGTTTCCCTTAGCGGATCAATTCGCGCCCTCTCCCTTCCCGCCCGCTT 1248  
Db 1202 CAGACCAACGGTTTCCCTTAGCGGATCAATTCG-----CCCCCCCCCT 1250  
QY 1249 AAGCTTACTGGCGAAGCGCTTGGATAGCGCGGTGCGTTGCTATATGTTATT 1308  
Db 1251 AAGCTTACTGGCGAAGCGCTTGGATAGCGCGGTGCGTTGCTATATGTTATT 1310  
QY 1309 TCCACATATTGCGCTTTTGGCAATGTGAGGCGCCGAAACCTGCGCCCTGCTTCTTTCG 1368  
Db 1311 TCCACATATTGCGCTTTTGGCAATGTGAGGCGCCGAAACCTGCGCCCTGCTTCTTTCG 1370  
QY 1369 ACAGCATTTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1428  
Db 1371 ACAGCATTTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1430  
QY 1429 GTGAAGGAAGCATCTCTGGAAGTCTTGAAGCAACACGCTCTGAGCGACCTT 1488  
Db 1431 GTGAAGGAAGCATCTCTGGAAGTCTTGAAGCAACACGCTCTGAGCGACCTT 1490  
QY 1489 TGAGGAGCGGGAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1548  
Db 1491 TGAGGAGCGGGAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1550  
QY 1549 TAAGATACACTGAAAGGCGGCAACCCCGACCTGCGACAGGTGCGACCTGTTGATGTTGTG 1608  
Db 1551 TAAGATACACTGAAAGGCGGCAACCCCGACCTGCGACAGGTGCGACCTGTTGATGTTGTG 1610  
QY 1609 GAAGAGTCAATGGCTCTCTCAAGGATTTCAAGAGGCTGAGGATGCCAGAG 1668  
Db 1611 GAAGAGTCAATGGCTCTCTCAAGGATTTCAAGAGGCTGAGGATGCCAGAG 1670  
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACTGCTTACATGTTTAG 1728  
Db 1671 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACTGCTTACATGTTTAG 1730

QY 1729 TCGAGGTTAAAAACGTCTAGGCCCCCGAAACACCGGGGACGTGGTTTCTTTGAAAAA 1788  
Db 1731 TCGAGGTTAAAAACGTCTAGGCCCCCGAAACACCGGGGACGTGGTTTCTTTGAAAAA 1790  
QY 1789 CAGGATAATACCATG----- 1803  
Db 1791 CAGGATAATACCATGACCGGGAGATGGCAGCATCTGTCGGAGCGGGTTTCGTAGGT 1850  
QY 1804 ----- 1803  
Db 1851 CTGATACTCTTGACCTTGTACCGCAGTATAAGCTTCTCTCGTAGGCTCATATGCTGG 1910  
QY 1804 ----- 1803  
Db 1911 TTACAATATTTTATCACAGGGCCGAGGCACACTTGAAGTGTGGATCCCGCCCTCAAC 1970  
QY 1804 ----- 1803  
Db 1971 GTTCGGGGGGCGGATGCGGTCTATCTCTCACGTGCGGATCCACCAGAGCTAATC 2030  
QY 1804 ----- 1803  
Db 2031 TTTACCATCACAAAATCTTGTCTGCCATCTCGTCCACTCATGTGTCTCCAGGCTGT 2090  
QY 1804 ----- 1803  
Db 2091 ATAACAAAAGTCCGTPACTTGTGCGCGCACACGGGCTCATTCGTGATGATGCTGGTG 2150  
QY 1804 ----- 1803  
Db 2151 CGGAAGTTGCTGGGGTCAATTATGTCCAAATGGTCTCATGAAGTTGGCGGACATGACA 2210  
QY 1804 ----- 1803  
Db 2211 GGTACGTACGTTTATGACCATCTCACCCACTCTGCGGACTGGGCCACCGCGCCCTACGA 2270  
QY 1804 ----- 1803  
Db 2271 GACCTTSGGTGGCAGTTTGAGCCCGTCTGCTTCTCTGATATGGAGACCAAGGTTATCACC 2330  
QY 1804 ----- 1803  
Db 2331 TGGGGGGGAGACACCGCGCGGTGTGGGGACATCATCTTGGGCTGCGCGCTCTCGGCCCGC 2390  
QY 1804 ----- 1803  
Db 2391 AGGGGAGGGAGATACATCTGGGACCGGACAGACAGCTTTGAAGGGCAGGGGTGGCGACTC 2450  
QY 1804 -----GCGCCTATTACGGCCTACTCCGAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
Db 2451 CTGCGCCCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCACT 2510  
QY 1861 AGCCTCACAGGCGGACAGGAACACAGTCTGAGGGGGAGGTCCAAAGTGTCTCCACGCA 1920  
Db 2511 AGCCTCACAGGCGGACAGGAACACAGTCTGAGGGGGAGGTCCAAAGTGTCTCCACGCA 2570  
QY 1921 ACACAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTTATCATGTGGC 1980  
Db 2571 ACACAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTTATCATGTGGC 2630  
QY 1981 GGTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCAATGTACCAATGTGGAC 2040  
Db 2631 GGTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCAATGTACCAATGTGGAC 2690  
QY 2041 CAGGACTCTGCTGGTGGCAAGCGCCCGCGGGCGGTTCCTTGAACCATCACTGCTC 2100  
Db 2691 CAGGACTCTGCTGGTGGCAAGCGCCCGCGGGCGGTTCCTTGAACCATCACTGCTC 2750  
QY 2101 GGCAGCTCGGACCTTTACTTGGTTCAGGAGCATGCGGATGTCTATCCGTTGCGCGCGG 2160  
Db 2751 GGCAGCTCGGACCTTTACTTGGTTCAGGAGCATGCGGATGTCTATCCGTTGCGCGCGG 2810

2161 GCGACAGAGGGGAGCTACTCTCCCGCCAGCGCCGCTCTCTACTTGAAGGCTCTTCG 2220 QY  
2811 GCGACAGAGGGGAGCTACTCTCCCGCCAGCGCCGCTCTCTACTTGAAGGCTCTTCG 2870 Db  
2221 GGGGTCACATGCTCTGCTCCCTCGGGCACTGCTGGGCACTTTTCGGGCTGCGGTGTC 2280 QY  
2871 GCGGTCACATGCTCTGCTCCCTCGGGCACTGCTGGGCACTTTTCGGGCTGCGGTGTC 2930 Db  
2281 ACCGAGGGTTCGAAGGGCTGACATTTGTACCGCTCGAGTCTATGGAACCACTATG 2340 QY  
2931 ACCGAGGGTTCGAAGGGCTGACATTTGTACCGCTCGAGTCTATGGAACCACTATG 2990 Db  
2341 CCGTCCCGGCTTTCACGCAAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGAGTG 2400 QY  
2991 CCGTCCCGGCTTTCAGGCAAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGAGTG 3050 Db  
2401 GGCATNTACAGCCCTTACTGTTAGGGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 2460 QY  
3051 GGCATNTACAGCCCTTACTGTTAGGGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 3110 Db  
2461 GCCAAGGGTATAAGGTGCTTCTGCTGAACCGCTCGCTCGCGCCACCTAGGTTTCGGG 2520 QY  
3111 GCCAAGGGTATAAGGTGCTTCTGCTGAACCGCTCGCTCGCGCCACCTAGGTTTCGGG 3170 Db  
2521 GCGTATGTTCTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGTAAAGCACTC 2580 QY  
3171 GCGTATGTTCTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGTAAAGCACTC 3230 Db  
2581 ACCAGGGTGCCCGCATCAGTACTTCCACTATGCGCAAGTTCCTTTCGCGAGCGGTGTC 2640 QY  
3231 ACCAGGGTGCCCGCATCAGTACTTCCACTATGCGCAAGTTCCTTTCGCGAGCGGTGTC 3290 Db  
2641 TCTGGGCGGCTATGACATCATATATGATGATGAGTGCCTCACTCACTGACTCGACCACT 2700 QY  
3291 TCTGGGCGGCTATGACATCATATATGATGATGAGTGCCTCACTCACTGACTCGACCACT 3350 Db  
2701 ATCTGGGCATCGGCACATGCTCTGCAAGCAAGCGAGAGCGGTGGAGCGCACTCGTCTGTG 2760 QY  
3351 ATCTGGGCATCGGCACATGCTCTGCAAGCAAGCGAGAGCGGTGGAGCGCACTCGTCTGTG 3410 Db  
2761 CTGCGACCGCTACGCTCCGGGATCGGTACCGTGCCACATTCACAAATCGAGAGGTG 2820 QY  
3411 CTGCGACCGCTACGCTCCGGGATCGGTACCGTGCCACATTCACAAATCGAGAGGTG 3470 Db  
2821 GCTCTGTGACGACTGGAGAAATCCCTTTTATGCAAGCAAGCCATCCCATCGAGACCATC 2880 QY  
3471 GCTCTGTGACGACTGGAGAAATCCCTTTTATGCAAGCAAGCCATCCCATCGAGACCATC 3530 Db  
2881 AAGGGGGGAGCACCTCATTTTCTGCCATTCCAGAAAGAAATGTGATGAGCTCGCCGCG 2940 QY  
3531 AAGGGGGGAGCACCTCATTTTCTGCCATTCCAGAAAGAAATGTGATGAGCTCGCCGCG 3590 Db  
2941 AAGCTGTCCGGCTCGGACTCAATGCTAGCATATTAACGGGGCTTGTATCGCTC 3000 QY  
3591 AAGCTGTCCGGCTCGGACTCAATGCTAGCATATTAACGGGGCTTGTATCGCTC 3650 Db  
3001 ATACCAATAGCGGAGAGCTCAATGCTAGCAACGGAGCGCTCTAATGACGGCTTTTACC 3060 QY  
3651 ATACCAATAGCGGAGAGCTCAATGCTAGCAACGGAGCGCTCTAATGACGGCTTTTACC 3710 Db  
3061 GCGATTTGCACTGATGATGCAATGCAATGATGTCACCCAGACAGTCTGACTCAGC 3120 QY  
3711 GCGATTTGCACTGATGATGCAATGCAATGATGTCACCCAGACAGTCTGACTCAGC 3770 Db  
3121 CTGACCCGAGCTTACCATTTGAGAGGAGCGCTGCGCAGAGCGCGGTGTCAGCTCG 3180 QY  
3771 CTGACCCGAGCTTACCATTTGAGAGGAGCGCTGCGCAGAGCGCGGTGTCAGCTCG 3830 Db  
3181 CAGCGCGAGGAGGAGCTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGCA 3240 QY  
3831 CAGCGCGAGGAGGAGCTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGCA 3890 Db  
3241 GAAAGGCGCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGGCTGT 3300 QY

3891 GAAAGGCGCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGGCTGT 3950 Db  
3301 GCTTGGTACGAGCTCACCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACTAAACACA 3360 QY  
3951 GCTTGGTACGAGCTCACCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACTAAACACA 4010 Db  
3361 CCAGGTTGCGCGCTGCGCAGGACCATCTGAGATTCTGGGAGAGCTTTTACAGGCTC 3420 QY  
4011 CCAGGTTGCGCGCTGCGCAGGACCATCTGAGATTCTGGGAGAGCTTTTACAGGCTC 4070 Db  
3421 ACCCACATAGAGCGCCCATTTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480 QY  
4071 ACCCACATAGAGCGCCCATTTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 4130 Db  
3481 CTGTAGCATACGAGCTACGCTGCGCAGGCTCAGGCTCAGCTCCTCATCTCTGTTGGAC 3540 QY  
4131 CTGTAGCATACGAGCTACGCTGCGCAGGCTCAGGCTCAGCTCCTCATCTCTGTTGGAC 4190 Db  
3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACGCGGCCAACGCCCTGCTG 3600 QY  
4191 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACGCGGCCAACGCCCTGCTG 4250 Db  
3601 TATAGGCTGGAGCGCTTCAAAACGAGTTTACTACACACACCCCATAAACCAATACATC 3660 QY  
4251 TATAGGCTGGAGCGCTTCAAAACGAGTTTACTACACACACCCCATAAACCAATACATC 4310 Db  
3661 ATGGCATCATGTCGGCTGACCTGGAGTGTGTCAGGACCTGAGTCTGCTGTTAGGCGGA 3720 QY  
4311 ATGGCATCATGTCGGCTGACCTGGAGTGTGTCAGGACCTGAGTCTGCTGTTAGGCGGA 4370 Db  
3721 GTCTAGCAGCTCTGGCGCGTATTGCTGCTGCAACAGGAGCGTGTCTTATGTTGGGAGG 3780 QY  
4371 GTCTAGCAGCTCTGGCGCGTATTGCTGCTGCAACAGGAGCGTGTCTTATGTTGGGAGG 4430 Db  
3781 ATCATCTTGTCCGAAAGCGCGCCATCATCTCCCGACAGGGAAGTCTTTACCGGGAGTTC 3840 QY  
4431 ATCATCTTGTCCGAAAGCGCGCCATCATCTCCCGACAGGGAAGTCTTTACCGGGAGTTC 4490 Db  
3841 GATGAGTGAAGAGTGGCGCTCAGCTCTCTTACATCGAACAGGGAATGAGCTCGCC 3900 QY  
4491 GATGAGTGAAGAGTGGCGCTCAGCTCTCTTACATCGAACAGGGAATGAGCTCGCC 4550 Db  
3901 GAAACAATCAACAGAGGCAATCGGTTGCTGCAACAGCACCAAGCAGGAGCGAGCT 3960 QY  
4551 GAAACAATCAACAGAGGCAATCGGTTGCTGCAACAGCACCAAGCAGGAGCGAGCT 4610 Db  
3961 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCTCTGAAAGCTTCTGCGGCGAAGCATATG 4020 QY  
4611 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCTCTGAAAGCTTCTGCGGCGAAGCATATG 4670 Db  
4021 TGGAAATTCATCAGGGGATCAATATTTAGAGGCTTGTGCACTCTGCTGCGGACCC 4080 QY  
4671 TGGAAATTCATCAGGGGATCAATATTTAGAGGCTTGTGCACTCTGCTGCGGACCC 4730 Db  
4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCGCGCTCACACCCACAT 4140 QY  
4731 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCGCGCTCACACCCACAT 4790 Db  
4141 ACCCTCTGTTTAAACATCTCGGGGATGAGTGGCGCGCCCAACTTGTCTCTCCAGCGCT 4200 QY  
4791 ACCCTCTGTTTAAACATCTCGGGGATGAGTGGCGCGCCCAACTTGTCTCTCCAGCGCT 4850 Db  
4201 GCTTCTGCTTTCGTTAGGCGCGCATCTGAGGCGGTGTGCGAGCATAGGCTTGGG 4260 QY  
4851 GCTTCTGCTTTCGTTAGGCGCGCATCTGAGGCGGTGTGCGAGCATAGGCTTGGG 4910 Db  
4261 AAGTGTCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTCGTGGCC 4320 QY  
4911 AAGTGTCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTCGTGGCC 4970 Db  
4321 TTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTCTCTGCT 4380 QY

Db 4971 TTTAAGGTATGAGCGGAGATGCCCTCCACGAGGACCTGGTTAACTACTCCCTGCT 5030  
QY 4381 ATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCGTGTGCCAGGATACCTGCGTGGCAC 4440  
Db 5031 ATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCGTGTGCCAGGATACCTGCGTGGCAC 5090  
QY 4441 GTGGGCCAGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGGTTTCGCTTGGCGG 4500  
Db 5091 GTGGGCCAGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGGTTTCGCTTGGCGG 5150  
QY 4501 GGTAAACCAAGTCTCCCCACGCACTATGTGCTGTGAGAGCGACGCTGCAGCACTGTCAC 4560  
Db 5151 GGTAAACCAAGTCTCCCCACGCACTATGTGCTGTGAGAGCGACGCTGCAGCACTGTCAC 5210  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTACAGTGTGTAAGAGGCTTCAAGTGGATCAAC 4620  
Db 5211 CAGATCTCTCTAGTCTTACCATCACTACAGTGTGTAAGAGGCTTCAAGTGGATCAAC 5270  
QY 4621 GAGGACTGTCTCCACGCCATCTCCGCTCGTGGCTAAGAGATGTTTGGGATTGGATATGC 4680  
Db 5271 GAGGACTGTCTCCACGCCATCTCCGCTCGTGGCTAAGAGATGTTTGGGATTGGATATGC 5330  
QY 4681 ACGGTGTGACTGATTCAAGACTGTGCTCCAGTCCAAAGCTCCGCGGATTCGCGGGA 4740  
Db 5331 ACGGTGTGACTGATTCAAGACTGTGCTCCAGTCCAAAGCTCCGCGGATTCGCGGGA 5390  
QY 4741 GTCCCTCTCTCTCATGTCAACGTGGTCAAGGGAGTCTGGCGGGGCGACGCACTATG 4800  
Db 5391 GTCCCTCTCTCTCATGTCAACGTGGTCAAGGGAGTCTGGCGGGGCGACGCACTATG 5450  
QY 4801 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAACGCTTCCATGAGG 4860  
Db 5451 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAACGCTTCCATGAGG 5510  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAACATTCGCCATTAACGGGTAC 4920  
Db 5511 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAACATTCGCCATTAACGGGTAC 5570  
QY 4921 ACCACGGGCTTGCAGGCTCCCGGCGCCAAATTTCTAGGGGCTGTGGCGGGTG 4980  
Db 5571 ACCACGGGCTTGCAGGCTCCCGGCGCCAAATTTCTAGGGGCTGTGGCGGGTG 5040  
QY 4981 GCTGCTGAGGATACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
Db 5631 GCTGCTGAGGATACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGGGCATG 5090  
QY 5041 ACCACTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
Db 5691 ACCACTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGCCCCCGAAATTTCTTCAAGAGTG 5150  
QY 5101 GATGGGTCGGTGTGACAGGTACGTCAGCGTGCAGGCTCCAAACCCCTCTACGGGAGGTC 5160  
Db 5751 GATGGGTCGGTGTGACAGGTACGTCAGCGTGCAGGCTCCAAACCCCTCTACGGGAGGTC 5220  
QY 5161 ACATTCCTGCTCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGTCGAGCCGGA 5220  
Db 5811 ACATTCCTGCTCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGTCGAGCCGGA 5280  
QY 5221 CCGGAGTACGATGTCTCACTTCCATGCTCACGACCCCTCCACATTAAGCGGAGAGC 5280  
Db 5871 CCGGAGTACGATGTCTCACTTCCATGCTCACGACCCCTCCACATTAAGCGGAGAGC 5280  
QY 5281 GCTAAGGTTAGGCTGGCAGGGATCTCCCTCTCTTGGCAGCTCATCAGTACGCGAG 5340  
Db 5931 GCTAAGGTTAGGCTGGCAGGGATCTCCCTCTCTTGGCAGCTCATCAGTACGCGAG 5390  
QY 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGATCCCGGAGCTGAC 5400  
Db 5991 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGTCACTGATCCCGGAGCTGAC 6050  
QY 5401 CTATCAGGSCCACTCTGTGGCGGAGAGATGGCGGGAACATCAACCCGCTGGAG 5460  
Db 6051 CTATCAGGSCCACTCTGTGGCGGAGAGATGGCGGGAACATCAACCCGCTGGAG 6110

QY 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAAAGCGGAGGAGTGA 5520  
Db 6111 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAAAGCGGAGGAGTGA 6170  
QY 5521 AGGGAAGTATCCCTTCCGCGGAGATCCTCGGAGGTTCAGGAAATTCCTCCAGGAGT 5580  
Db 6171 AGGGAAGTATCCCTTCCGCGGAGATCCTCGGAGGTTCAGGAAATTCCTCCAGGAGT 6230  
QY 5581 CCATATGGGACCGCCCGGATTAACAACCTCTCACTGTAGAGTCTTGGAAAGGACCCGAC 5640  
Db 6231 CCATATGGGACCGCCCGGATTAACAACCTCTCACTGTAGAGTCTTGGAAAGGACCCGAC 6290  
QY 5641 TAGCTCTCTCACTGATGATACAGGCTGTCAATTCGCGCTGCGAAGGCCCTCCGATACCA 5700  
Db 6291 TAGCTCTCTCACTGATGATACAGGCTGTCAATTCGCGCTGCGAAGGCCCTCCGATACCA 6350  
QY 5701 CTTCCACGGAGGAGGAGGTTGCTGTCTAGAAATCTCGTCCGCTCGACACGCGCACG 5760  
Db 6351 CTTCCACGGAGGAGGAGGTTGCTGTCTAGAAATCTCGTCCGCTCGACACGCGCACG 6410  
QY 5761 GAGCTCGCCACAAAGACCTTTCGGCAGCTCCGAATCTCGTCCGCTCGACACGCGCACG 5820  
Db 6411 GAGCTCGCCACAAAGACCTTTCGGCAGCTCCGAATCTCGTCCGCTCGACACGCGCACG 6470  
QY 5821 ACGGCTCTCTGACAGAGCTTCGAGAGCGGAGCGGAGATCCGAGTGTGATGCTGATC 5880  
Db 6471 ACGGCTCTCTGACAGAGCTTCGAGAGCGGAGCGGAGATCCGAGTGTGATGCTGATC 6530  
QY 5881 TCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG 5940  
Db 6531 TCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG 6590  
QY 5941 TCTACCGTAAAGGAGGAGCTAGTGAAGAGCTGCTGCTGCTGCTGCTGCTTACATG 6000  
Db 6591 TCTACCGTAAAGGAGGAGCTAGTGAAGAGCTGCTGCTGCTGCTGCTGCTTACATG 6050  
QY 6001 ACAGGCGCTGATCAGCGATGCGTGGAGGAAACCAAGTGCCTCAATCAATGACATG 6060  
Db 6651 ACAGGCGCTGATCAGCGATGCGTGGAGGAAACCAAGTGCCTCAATCAATGACATG 6120  
QY 6061 AGCAACTCTTGTCTGCTCACCACAACTTGTCTATGTACACATCTCGACGCGAAG 6120  
Db 6711 AGCAACTCTTGTCTGCTCACCACAACTTGTCTATGTCTAACAATCTCGACGCGAAG 6170  
QY 6121 CTGCGGACAGAAAGTCACTTTGACAGCTGAGGTCCTTGGACGACCACTACCGGAC 6180  
Db 6771 CTGCGGACAGAAAGTCACTTTGACAGCTGAGGTCCTTGGACGACCACTACCGGAC 6240  
QY 6181 GTGCTCAAGAGATGAAGCGGAGGCTGCAAGTTAAGCTTAACTCTATCCGTGGAG 6240  
Db 6831 GTGCTCAAGAGATGAAGCGGAGGCTGCAAGTTAAGCTTAACTCTATCCGTGGAG 6290  
QY 6241 GAAGCTGTAAAGTACGCCCCACATTCGCGAGATCTAAATTTGGCTATGGGCAAG 6300  
Db 6891 GAAGCTGTAAAGTACGCCCCACATTCGCGAGATCTAAATTTGGCTATGGGCAAG 6350  
QY 6301 GAGCTCGGAACTATTCAGCAAGGCGTTAAACCACTCCGCTCCGTTGGAGGAGCTTGG 6360  
Db 6951 GAGCTCGGAACTATTCAGCAAGGCGTTAAACCACTCCGCTCCGTTGGAGGAGCTTGG 7010  
QY 6361 CTGGAAGACACTGAGACACCAATTCACCACTATGCGCAAAATGAGGTTTCTGC 6420  
Db 7011 CTGGAAGACACTGAGACACCAATTCACCACTATGCGCAAAATGAGGTTTCTGC 7070  
QY 6421 GTCCACACAGAGAAAGGGGCGGAGCGAGCTCGCTTATCGTATTCACGATTTGGG 6480  
Db 7071 GTCCACACAGAGAAAGGGGCGGAGCGAGCTCGCTTATCGTATTCACGATTTGGG 7130  
QY 6481 GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCTCCCTCAGGCGGT 6540  
Db 7131 GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCTCCCTCAGGCGGT 7190





Db 242 CGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCCCTTGTGTAATGCTGCTGATAGG 301  
QY 302 TGCTTGGAGTGGCCCGGAGGTCTCTAGACCGTGCACCATGAGCAGAAATCCTAAACC 361  
Db 302 TGCTTGGAGTGGCCCGGAGGTCTCTAGACCGTGCACCATGAGCAGAAATCCTAAACC 361  
QY 362 TCAAGAAACCAAGGCGCGCCATGATTGAACAAGATGATGACCGAGTTCACCGAGTTCCTC 421  
Db 362 TCAAGAAACCAAGGCGCGCCATGATTGAACAAGATGATGACCGAGTTCACCGAGTTCCTC 421  
QY 422 GGCCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAAATCGGCTGCTC 481  
Db 422 GGCCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAAATCGGCTGCTC 481  
QY 482 TGATGCGCGCTGTTCCGGCTGTACGGCAGGGGCGCCGGTCTTTTGTCAAGACCGA 541  
Db 482 TGATGCGCGCTGTTCCGGCTGTACGGCAGGGGCGCCGGTCTTTTGTCAAGACCGA 541  
QY 542 CTTGTCGGTGCCTGAATGAATGACAGGACGAGGCGCGGTATCGTGGCTGCGCAC 601  
Db 542 CTTGTCGGTGCCTGAATGAATGACAGGACGAGGCGCGGTATCGTGGCTGCGCAC 601  
QY 602 GACGGCGTTCCTTGGCAGCTGTGCTCGACGTGTCTCACTGAAGCGGGAAGGACTGGCT 661  
Db 602 GACGGCGTTCCTTGGCAGCTGTGCTCGACGTGTCTCACTGAAGCGGGAAGGACTGGCT 661  
QY 662 GCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTCACTCACTGCTCTCTGCGGAGAA 721  
Db 662 GCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTCACTCACTGCTCTCTGCGGAGAA 721  
QY 722 AGTATCATCATGCTGATGATCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCC 781  
Db 722 AGTATCATCATGCTGATGATCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCC 781  
QY 782 ATTTCGACCAAGCGAAACATCGCTCGAGCGAGCACGTACTCGGATGGAAGCGGTCT 841  
Db 782 ATTTCGACCAAGCGAAACATCGCTCGAGCGAGCACGTACTCGGATGGAAGCGGTCT 841  
QY 842 TGTGATCAGATGATCTGCAAGAGATCAGGGGCTCGGCGCACCGAACTGTTGCG 901  
Db 842 TGTGATCAGATGATCTGCAAGAGATCAGGGGCTCGGCGCACCGAACTGTTGCG 901  
QY 902 CAGGCTCAAGCGCGATGCGGACGCGAGGATCTGCTGACCATGCGATGCGCTG 961  
Db 902 CAGGCTCAAGCGCGATGCGGACGCGAGGATCTGCTGACCATGCGATGCGCTG 961  
QY 962 CTTGCCGAATATCATGTTGGAATGCGCGCTTTTCTGGATTATCGACTGTGCGCGCT 1021  
Db 962 CTTGCCGAATATCATGTTGGAATGCGCGCTTTTCTGGATTATCGACTGTGCGCGCT 1021  
QY 1022 GGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCGGTGATGCTGAAAGCT 1081  
Db 1022 GGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCGGTGATGCTGAAAGCT 1081  
QY 1082 TGGCGCGAATGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTGCA 1141  
Db 1082 TGGCGCGAATGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTGCA 1141  
QY 1142 GGCATCGCTTCTATCGCTTCTTGAAGAGTTCCTTCTGAGTT-----TAAAC 1189  
Db 1142 GGCATCGCTTCTATCGCTTCTTGAAGAGTTCCTTCTGAGTTTCGCGCCACAGATTAAAC 1201  
QY 1190 AGACCAACAGGTTTCCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTA 1249  
Db 1202 AGACCAACAGGTTTCCTTAGCGGATCAATTCGCGCCCTA-----CCCGCCCGCTA 1250  
QY 1250 ACCTTACTGCGGACCGCTTGAATAAGCGCGGTGCTGTTTGTCTATGTTATTTT 1309  
Db 1251 ACCTTACTGCGGACCGCTTGAATAAGCGCGGTGCTGTTTGTCTATGTTATTTT 1310  
QY 1310 CCACCATATTGCGCTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTCTCTTGA 1369

Db 1311 CCACCATATTGCGCTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGCTCTTCTTGA 1370  
QY 1370 CGAGACTTCCTAGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1429  
Db 1371 CGAGACTTCCTAGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1430  
QY 1430 TGAAGNAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAAACAACTGTTGAGGACCCCTT 1489  
Db 1431 TGAAGNAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAAACAACTGTTGAGGACCCCTT 1490  
QY 1490 GCAGGACGCGAAACCCGCCACCTGGCGACAGGTGCTCTCGGCCAAAAGCACGTGTAT 1549  
Db 1491 GCAGGACGCGAAACCCGCCACCTGGCGACAGGTGCTCTCGGCCAAAAGCACGTGTAT 1550  
QY 1550 AAGATACCTCTGCAAGGCGGCAACCCAGTGCCACGTTGTGAGTTGATGTTGTTAGT 1609  
Db 1551 AAGATACCTCTGCAAGGCGGCAACCCAGTGCCACGTTGTGAGTTGATGTTGTTAGT 1610  
QY 1610 AAGAGTCAAAATGGCTCTCAAGCGTATTCACAGGGGCTGAAGGATGCCAGAAAG 1669  
Db 1611 AAGAGTCAAAATGGCTCTCAAGCGTATTCACAGGGGCTGAAGGATGCCAGAAAG 1670  
QY 1670 TACCCATTCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGTAGT 1729  
Db 1671 TACCCATTCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGTAGT 1730  
QY 1730 CGAGGTTAAAAACGCTTAGGCCCCCGAAACCAAGGGGACGTGTTTCTTTGAAAAAC 1789  
Db 1731 CGAGGTTAAAAACGCTTAGGCCCCCGAAACCAAGGGGACGTGTTTCTTTGAAAAAC 1790  
QY 1790 ACGATTAATACG----- 1803  
Db 1791 ACGATTAATACGACCGGAGATGCGACATCGTCCGAGGCGCGTTCCTTAGGTC 1850  
QY 1804 ----- 1803  
Db 1851 TGATACTCTTGACCTTGTACCGCACTATAAGCTGTTCTCGTAGGCTCATATGTTGGT 1910  
QY 1804 ----- 1803  
Db 1911 TACAATATTTATACACAGGCGGACACACTTGCAGAGTGTGATCCTCCCGCTCAACG 1970  
QY 1804 ----- 1803  
Db 1971 TTCGGGGGCGCGAGTCCGTCATCTCCTCACGTGCGGATCCACCCAGAGCTAATCT 2030  
QY 1804 ----- 1803  
Db 2031 TTACCATCACCATAATCTTGCTCGCCATCTCGGTCCACTCATGTTGCTCCAGGCTGTA 2090  
QY 1804 ----- 1803  
Db 2091 TAACCAAGTGGCTACTCTGTCGCGCACACGCGGCTCATCTGTCATGTCATGTCGTGTC 2150  
QY 1804 ----- 1803  
Db 2151 GGAAGTGTGTCGGGTCAATATGTCCAAATGCTCTCATGAAGTTGGCGCATGACAG 2210  
QY 1804 ----- 1803  
Db 2211 GTACGTACGTTTATGACCATCTCACCCACTGGGACTGGGCCACGCGGCTTAGGAG 2270  
QY 1804 ----- 1803  
Db 2271 ACCTTGGGTGGCAGTTGAGCCGCTGCTCTCTGATATGAGACCAAGTTATCACCT 2330  
QY 1804 ----- 1803  
Db 2331 GGGGGGACACCGCGGCTGTGGGACATCATCTTGGGCTGCGCGCTCTCGCGCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGGAGGAGATACATCTGGGACCGGACAGACCTTTGAAGGGGAGGGGTGCGGACTCC 2450

QY	1804	--GCGCCTTTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACTA	1861	QY	2882	AGGGGGGAGGACCTCTTTCTCCCATTTCCAGAGAAATGTGATGAGCTCGCGCGA	2941
DB	2451	TGCGCCCTATTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACTA	2510	DB	3531	AGGGGGGAGGACCTCTTTCTCCCATTTCCAGAGAAATGTGATGAGCTCGCGCGA	3590
QY	1862	GCCTCACAGGCGCGGACAGGAACAGGTTCGAGGGGAGGTCAAGTGGTCTCCACCGAA	1921	QY	2942	AGCTTCGGGCTCGGACTCAATGCTAGCATATTACCGGGCCTTGATGATCGGTCA	3001
DB	2511	GCCTCACAGGCGCGGACAGGAACAGGTTCGAGGGGAGGTCAAGTGGTCTCCACCGAA	2570	DB	3591	AGCTTCGGGCTCGGACTCAATGCTAGCATATTACCGGGCCTTGATGATCGGTCA	3650
QY	1922	CACAACTTTCTGTGGCACTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGGCG	1981	QY	3002	TACCAACTAGCGGAGACGTCAATGCTAGTACCAACGAGCTCTAATGAGGGCTTTACCG	3061
DB	2571	CACAACTTTCTGTGGCACTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGGCG	2630	DB	3651	TACCAACTAGCGGAGACGTCAATGCTAGTACCAACGAGCTCTAATGAGGGCTTTACCG	3710
QY	1982	GCTCAAGACCTTCCCGGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC	2041	QY	3062	GGATTTCCGACTCAGTGAATCGACTGCAATACATGTGTCAACAGACAGTCACTTCAGCC	3121
DB	2631	GCTCAAGACCTTCCCGGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC	2690	DB	3711	GGATTTCCGACTCAGTGAATCGACTGCAATACATGTGTCAACAGACAGTCACTTCAGCC	3770
QY	2042	AGGACCTCGTCCGGCTGGCAAGCGCCCCCGGGCGCGTTCTTTGACACCATGCACTGG	2101	QY	3122	TGGACCCCACTTCACCAATTGAGACGACCGCTGCCAAGACGCGGTGTCACTTCGCG	3181
DB	2691	AGGACCTCGTCCGGCTGGCAAGCGCCCCCGGGCGCGTTCTTTGACACCATGCACTGG	2750	DB	3771	TGGACCCCACTTCACCAATTGAGACGACCGCTGCCAAGACGCGGTGTCACTTCGCG	3830
QY	2102	GCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCGGATGTCAATCCGGTGCACCGCGGG	2161	QY	3182	AGCGGAGGACGAGACTGGTAGGGGAGGATGGGCAATTTACAGTGTGTGACTCCAGGAG	3241
DB	2751	GCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCGGATGTCAATCCGGTGCACCGCGGG	2810	DB	3831	AGCGGAGGACGAGACTGGTAGGGGAGGATGGGCAATTTACAGTGTGTGACTCCAGGAG	3890
QY	2162	GGCAGACAGGGGGAGCCTACTCTCCCGCAGCGCGTCTCTACTTGAAGGGCTCTTCGG	2221	QY	3242	AACGGCCCTCGGCGCATGTTCCGATTCCTCGGTCTCTGCGAGTGTATGACGCGGGGTGTG	3301
DB	2811	GGCAGACAGGGGGAGCCTACTCTCCCGCAGCGCGTCTCTACTTGAAGGGCTCTTCGG	2870	DB	3891	AACGGCCCTCGGCGCATGTTCCGATTCCTCGGTCTCTGCGAGTGTATGACGCGGGGTGTG	3950
QY	2222	GGGCTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCACTTTTGGGCTGCCGTGTGCA	2281	QY	3302	CTTGTAGAGCTACGCGCCCGGAGACTCAAGTTAGTGTGGGTGTACTTAACACACAC	3361
DB	2871	GGGCTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCACTTTTGGGCTGCCGTGTGCA	2930	DB	3951	CTTGTAGAGCTACGCGCCCGGAGACTCAAGTTAGTGTGGGTGTACTTAACACACAC	4010
QY	2282	CCCGAGGGGTTCGGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACACATATGC	2341	QY	3362	GAGGCTTCCCGCTCTGCCAGGACCATCTGGAGTCTGGGAGAGCTCTTTACAGGCTCA	3421
DB	2931	CCCGAGGGGTTCGGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACACATATGC	2990	DB	4011	CAGGCTTCCCGCTCTGCCAGGACCATCTGGAGTCTGGGAGAGCTCTTTACAGGCTCA	4070
QY	2342	GGTCCCGGTCTTCAGGCAACTCGTCCCGCTCCGCGCTACCGGAGACTTCGAGGTGG	2401	QY	3422	CCCATATAGACGCCCATTTCTTGTCCAGACTAAGCAGGACGAGACACTTCCCTTACC	3481
DB	2991	GGTCCCGGTCTTCAGGCAACTCGTCCCGCTCCGCGCTACCGGAGACTTCGAGGTGG	3050	DB	4071	CCCATATAGACGCCCATTTCTTGTCCAGACTAAGCAGGACGAGACACTTCCCTTACC	4130
QY	2402	CCCATCTACACGCCCTACTGTAGCGCAAGAGCACTAAGGTGCGGCTGGGTATGCA	2461	QY	3482	TGCTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTTCCATCGTGGAGC	3541
DB	3051	CCCATCTACACGCCCTACTGTAGCGCAAGAGCACTAAGGTGCGGCTGGGTATGCA	3110	DB	4131	TGCTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTTCCATCGTGGAGC	4190
QY	2462	CCCAAGGTATAGGTGCTTGTCTGAAACCGCTCCGCTCCGCGCCACCCCTAGGTTCGGG	2521	QY	3542	AAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCCAACGCCCTCTCT	3601
DB	3111	CCCAAGGTATAGGTGCTTGTCTGAAACCGCTCCGCTCCGCGCCACCCCTAGGTTCGGG	3170	DB	4191	AAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCCAACGCCCTCTCT	4250
QY	2522	CGTATATGTCTAAGGCACATGTTATCGACCTTAACATCAGAACCGGGGTAAGACCATCA	2581	QY	3602	ATAGGCTGGGAGCCGTTCAAAACGAGGTCTACTACCAACACCCCATTAACCAATACATCA	3661
DB	3171	CGTATATGTCTAAGGCACATGTTATCGACCTTAACATCAGAACCGGGGTAAGACCATCA	3230	DB	4251	ATAGGCTGGGAGCCGTTCAAAACGAGGTCTACTACCAACACCCCATTAACCAATACATCA	4310
QY	2582	CCACGGGTCCCGCCATCACTCCACCTATGGCAAGTTTCTTGCCGACGGTGGTGTCT	2641	QY	3662	TGGCATATGTCGGCTGACCTGAGGTGTGTCAAGGACCTTGGGTGTGTGAGGCGGAG	3721
DB	3231	CCACGGGTCCCGCCATCACTCCACCTATGGCAAGTTTCTTGCCGACGGTGGTGTCT	3290	DB	4311	TGGCATATGTCGGCTGACCTGAGGTGTGTCAAGGACCTTGGGTGTGTGAGGCGGAG	4370
QY	2642	CTGGGGGCGCCTATGACATCAATAATGTATGATGAGTGCCACTCAACTGACCACTA	2701	QY	3722	TCCTAGAGCTCTGCGCGGTATTCCTGACAAACAGGACGCTGTGTCATTGTGGGAGGA	3781
DB	3291	CTGGGGGCGCCTATGACATCAATAATGTATGATGAGTGCCACTCAACTGACCACTA	3350	DB	4371	TCCTAGAGCTCTGCGCGGTATTCCTGACAAACAGGACGCTGTGTCATTGTGGGAGGA	4430
QY	2702	TCCTGGGATCGGCACAGTCTTGGACCAAGCGGAGACGCTGGAGCGCACTCTGCTGC	2761	QY	3782	TCATCTTGTCCGAAAGCCGCGCATCAATTCGCGACAGGGAAGTCTTTTACCGGGAGTTCG	3841
DB	3351	TCCTGGGATCGGCACAGTCTTGGACCAAGCGGAGACGCTGGAGCGCACTCTGCTGC	3410	DB	4431	TCATCTTGTCCGAAAGCCGCGCATCAATTCGCGACAGGGAAGTCTTTTACCGGGAGTTCG	4490
QY	2762	TGCGCACCGCTACGCTCCGGGATCGGTTCACCGTGCCACATCCAAACATCGAGGAGTGG	2821	QY	3842	ATGAGATGGAAGAGTGTGCGCTCACCTCCCTTATCATGAAACAGGGAATGCACTGCGCG	3901
DB	3411	TGCGCACCGCTACGCTCCGGGATCGGTTCACCGTGCCACATCCAAACATCGAGGAGTGG	3470	DB	4491	ATGAGATGGAAGAGTGTGCGCTCACCTCCCTTATCATGAAACAGGGAATGCACTGCGCG	4550
QY	2822	CTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGGCATCCCGCATCGAGACCATCA	2881	QY	3902	AAACATTCAAAACAGAGGCAATCGGTTCTGCAAAACAGCCACCAAGCAAGCGAGGCTG	3961
DB	3471	CTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGGCATCCCGCATCGAGACCATCA	3530	DB	4551	AAACATTCAAAACAGAGGCAATCGGTTCTGCAAAACAGCCACCAAGCAAGCGAGGCTG	4610
QY				QY	3962	CTGCTCCGCTGTGGAAATCCAAAGTGGCGGACCCCTCGAAAGCCCTTCTGGGGAAGCATATGT	4021

4611 DB ||||| CTGCTCCGCTGGTGGAAATCCAGTGGCGGACCTCCAGACCTTCTGGGGAACATATGT 4670  
4022 QY GGAATTTTCATCAGCGGATACAATATTTAGCAGCTTGTCCACTCTGCCTGGCAACCCCG 4081  
4671 DB GGAATTTTCATCAGCGGATACAATATTTAGCAGCTTGTCCACTCTGCCTGGCAACCCCG 4730  
4082 QY CGATAGCATCATGTATGGGATTCACAGCTCTATACACCGCCGCTCACACCCCAACATA 4141  
4731 DB CGATAGCATCATGTATGGGATTCACAGCTCTATACACCGCCGCTCACACCCCAACATA 4790  
4142 QY CCTCTCTGTTTAACTCTCGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCTG 4201  
4791 DB CCTCTCTGTTTAACTCTCGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCTG 4850  
4202 QY CTTCTCTGTTTCGAGCGCGGATCGCTGGAGCGCTGTGGCAGCATAGGCTTTGGGA 4261  
4851 DB CTTCTCTGTTTCGAGCGCGGATCGCTGGAGCGCTGTGGCAGCATAGGCTTTGGGA 4910  
4262 QY AGTGCTTGTGGATATTTGGCAGGTATGGACAGGGGTGGCAGGGGCTCGTGCCT 4321  
4911 DB AGTGCTTGTGGATATTTGGCAGGTATGGACAGGGGTGGCAGGGGCTCGTGCCT 4970  
4322 QY TTAAGGTCATGACGGCGAGATGCCCTCCACCGAGACCTGTGTTAACTACTCCCTGCTA 4381  
4971 DB TTAAGGTCATGACGGCGAGATGCCCTCCACCGAGACCTGTGTTAACTACTCCCTGCTA 5030  
4382 QY TCCTCTCCCTCGCGCCTTAGTCTGGGGTCTGTGGCGAGCATCTGCTCGGCAGC 4441  
5031 DB TCCTCTCCCTCGCGCCTTAGTCTGGGGTCTGTGGCGAGCATCTGCTCGGCAGC 5090  
4442 QY TGGGCCCCAGGGAGGGGCTGTGCTGCTGAGTGAACCGGCTGATAGCTTCGCTTCGGGG 4501  
5091 DB TGGGCCCCAGGGAGGGGCTGTGCTGCTGAGTGAACCGGCTGATAGCTTCGCTTCGGGG 5150  
4502 QY GTAACACAGCTCTCCCCACCACTATGCTGCTCAGAGCGACGCTGCAGCAGCTGTCACTC 4561  
5151 DB GTAACACAGCTCTCCCCACCACTATGCTGCTCAGAGCGACGCTGCAGCAGCTGTCACTC 5210  
4562 QY AGATCTCTCTAGTCTTACCATCACTCAGCTGTGAAGAGGCTTCACCAAGTGGATCAACG 4621  
5211 DB AGATCTCTCTAGTCTTACCATCACTCAGCTGTGAAGAGGCTTCACCAAGTGGATCAACG 5270  
4622 QY AGGACTGCTCCAGCCATGCTCCGGCTGTGGCTAGAGATGTTGGATGGATATGCA 4681  
5271 DB AGGACTGCTCCAGCCATGCTCCGGCTGTGGCTAGAGATGTTGGATGGATATGCA 5330  
4682 QY CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTGCGCGATTCGCCGGAG 4741  
5331 DB CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTGCGCGATTCGCCGGAG 5390  
4742 QY TCCCTCTCTTCTCATGTCAAAGTGGGTPAACAGGAGTCTGGCGGGGCGACGGCATATGC 4801  
5391 DB TCCCTCTCTTCTCATGTCAAAGTGGGTPAACAGGAGTCTGGCGGGGCGACGGCATATGC 5450  
4802 QY AATACGACTGCCATGTGGAGCAGATCACCGGACATGTGAATAACGGTTCCATGAGGA 4861  
5451 DB AATACGACTGCCATGTGGAGCAGATCACCGGACATGTGAATAACGGTTCCATGAGGA 5510  
4862 QY TCGTGGGGCTTAGACTGTAGTAAACAGCTGGATGGAACATTTCCCATTAACCGCTACA 4921  
5511 DB TCGTGGGGCTTAGACTGTAGTAAACAGCTGGATGGAACATTTCCCATTAACCGCTACA 5570  
4922 QY CCAAGGCCCCCTGACGCCCTCCCGGCGCAAAATTTATCTAGGGGCTGTGGCGGTGG 4981  
5571 DB CCAAGGCCCCCTGACGCCCTCCCGGCGCAAAATTTATCTAGGGGCTGTGGCGGTGG 5630  
4982 QY CTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGATTTCCACTAGTACGGGCAATGA 5041  
5631 DB CTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGATTTCCACTAGTACGGGCAATGA 5690  
5042 QY CCACTGACACGTAAGTGGCGGTGTCAGGTTCCGGCCCCCGAATCTTTCACAGAGTGG 5101

5691 DB CCACTGACACGTAAGTGGCGGTGTCAGGTTCCGGCCCCCGAATTTCTTACAGAAAGTGG 5750  
5102 QY ATGGGGTGGGTTGCAAGGTACGCTCCAGCGTGGAAAACCCCTCTCTACGGGAGGAGTCA 5161  
5751 DB ATGGGGTGGGTTGCAAGGTACGCTCCAGCGTGGAAAACCCCTCTCTACGGGAGGAGTCA 5810  
5162 QY CATTCCTGTTCCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGTCGAGCCCGAAC 5221  
5811 DB CATTCCTGTTCCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGTCGAGCCCGAAC 5870  
5222 QY CGGACGTAGCAGTGTCACTTCCATGCTCACGACCCCTCCCAATTTACGCGGAGACGG 5281  
5871 DB CGGACGTAGCAGTGTCACTTCCATGCTCACGACCCCTCCCAATTTACGCGGAGACGG 5930  
5282 QY CTAAGGTTAGGTCGCGCAGGGATCTCCCCCTCTCTTGGCGAGCTCATGCTAGCCAGC 5341  
5931 DB CTAAGGTTAGGTCGCGCAGGGATCTCCCCCTCTCTTGGCGAGCTCATGCTAGCCAGC 5990  
5342 QY TGTCTCGGCTTCTTGAAGGCAACATGCATACCGTGCATGCTCCCGGACGCTGACC 5401  
5991 DB TGTCTCGGCTTCTTGAAGGCAACATGCATACCGTGCATGCTCCCGGACGCTGACC 6050  
5402 QY TCATCGAGGCCAACCTCTCTGTCGCGCAGGAGATGGCGGGAACATCACCCGCTGGAGT 5461  
6051 DB TCATCGAGGCCAACCTCTCTGTCGCGCAGGAGATGGCGGGAACATCACCCGCTGGAGT 6110  
5462 QY CAGAAATTAAGGTAGTAAATTTTGGACTCTTTGAGCGGCTCCAGCGGAGGAGATGAGA 5521  
6111 DB CAGAAATTAAGGTAGTAAATTTTGGACTCTTTGAGCGGCTCCAGCGGAGGAGGATGAGA 6170  
5522 QY GGAAGTATTCGCTTCGCGGAGATCTCGGAGGTCCAGAAATTCCTTCGAGCGATGC 5581  
6171 DB GGAAGTATTCGCTTCGCGGAGATCTCGGAGGTCCAGAAATTCCTTCGAGCGATGC 6230  
5582 QY CCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCCCTGGAAGGACCCGACT 5641  
6231 DB CCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCCCTGGAAGGACCCGACT 6290  
5642 QY AGTCTCCTCAGTGTACACGGGTGCCATGTCGCTCCAGAGGCCCTCCGATACCAC 5701  
6291 DB AGTCTCCTCAGTGTACACGGGTGCCATGTCGCTCCAGAGGCCCTCCGATACCAC 6350  
5702 QY CTCACGGGAGGAGGAGCGGTGCTCTCGAATCTACCGTGTCTTCGCTTCGCGG 5761  
6351 DB CTCACGGGAGGAGGAGCGGTGCTCTCGAATCTACCGTGTCTTCGCTTCGCGG 6410  
5762 QY AGTCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCCGCGTCCGACGCGCACGGCAA 5821  
6411 DB AGTCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCCGCGTCCGACGCGCACGGCAA 6470  
5822 QY CGGCTCTCTGTACCAAGCCCTCCGACGACGGCGAGGATCCGAGTTGAGTCTGACT 5881  
6471 DB CGGCTCTCTGTACCAAGCCCTCCGACGACGGCGAGGATCCGAGTTGAGTCTGACT 6530  
5882 QY CTTCCATGCCCTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGTCTTGGT 5941  
6531 DB CTTCCATGCCCTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGTCTTGGT 6590  
5942 QY CTAACGTAAGCGAGGAGGCTAGTGGAGCTGCTGCTGCTCGATGTCTACACATGGA 6001  
6591 DB CTAACGTAAGCGAGGAGGCTAGTGGAGCTGCTGCTGCTCGATGTCTACACATGGA 6650  
6002 QY CAGGCGCCTGTATCAGCCATGCGGTGCGGAGGAAACCAAGCTGCCATCAATGCACTGA 6061  
6651 DB CAGGCGCCTGTATCAGCCATGCGGTGCGGAGGAAACCAAGCTGCCATCAATGCACTGA 6710  
6062 QY GCAACTCTTTGCTCCCTCACCACCACTTGTCTATCTCAACATCTCGCAGCGCAAGCC 6121  
6711 DB GCAACTCTTTGCTCCCTCACCACCACTTGTCTATCTCAACATCTCGCAGCGCAAGCC 6770  
6122 QY TGGCGCAGAAAGGTACCTTTGACAGACTGAGGTCCTGGACGACCACTACCGGGAGC 6181  
6771 DB TGGCGCAGAAAGGTACCTTTGACAGACTGAGGTCCTGGACGACCACTACCGGGAGC 6830

QY 6182 TGCTCAAGGAGATGAAGCGGAGCGCTCCACAGTTAAAGCTTAAACTTCTATCCGTGGAGG 6241  
Db 6831 TGCTCAAGGAGATGAAGCGGAGCGCTCCACAGTTAAAGCTTAAACTTCTATCCGTGGAGG 6890  
QY 6242 AAGCCTGTAAAGTGAAGCGGCGCCACACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAAAGG 6301  
Db 6891 AAGCCTGTAAAGTGAAGCGGCGCCACACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAAAGG 6950  
QY 6302 AGCTCCGGAACTATCCAGCAAGCGCGTTAAACCAATCCGCTCCGTGTGGAAAGCACTTGC 6361  
Db 6951 AGCTCCGGAACTATCCAGCAAGCGCGTTAAACCAATCCGCTCCGTGTGGAAAGCACTTGC 7010  
QY 6362 TGGAGACACTGAGACACCAATTTGACACACCATCATGSCAAAATAGAGTTTCTCGG 6421  
Db 7011 TGGAGACACTGAGACACCAATTTGACACACCATCATGSCAAAATAGAGTTTCTCGG 7070  
QY 6422 TCCAAACGAGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTCAGATTTGGGG 6481  
Db 7071 TCCAAACGAGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTCAGATTTGGGG 7130  
QY 6482 TTCGTGTGCGAGAAATGGCCCTTTAGGATGCTCTCACCTCCCTCGAGCCGTGA 6541  
Db 7131 TTCGTGTGCGAGAAATGGCCCTTTAGGATGCTCTCACCTCCCTCGAGCCGTGA 7190  
QY 6542 TGGGCTCTTATACGGATTCCAAATACTCTCTGGACAGCGGTGAGTTCCCTGGTGAATG 6601  
Db 7191 TGGGCTCTTATACGGATTCCAAATACTCTCTGGACAGCGGTGAGTTCCCTGGTGAATG 7250  
QY 6602 COTGAAAGCGAAGAAATGCCCTATGGCTTGGCAATATGACACCGCTGTTTTGACTCAA 6661  
Db 7251 COTGAAAGCGAAGAAATGCCCTATGGCTTGGCAATATGACACCGCTGTTTTGACTCAA 7310  
QY 6662 CGGTCACTGAGAAATGACATCGTGTGAGAGTCAATCTACCAATGTTGACTGGGCC 6721  
Db 7311 CGGTCACTGAGAAATGACATCGTGTGAGAGTCAATCTACCAATGTTGACTGGGCC 7370  
QY 6722 CCGAAGCGAGACAGGCCATAAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTGA 6781  
Db 7371 CCGAAGCGAGACAGGCCATAAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTGA 7430  
QY 6782 CTAATTTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGA 6841  
Db 7431 CTAATTTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGA 7490  
QY 6842 CCAGTCCGGTAAATACCTTACATGTTACTTGAAGCGCGCTGCGGCCCTGTGAGCTGCGA 6901  
Db 7491 CCAGTCCGGTAAATACCTTACATGTTACTTGAAGCGCGCTGCGGCCCTGTGAGCTGCGA 7550  
QY 6902 AGCTCCAGGACTGCACGATGCTCGTATCGGAGACGACCTTGTCTTATCTGTGAAAGCG 6961  
Db 7551 AGCTCCAGGACTGCACGATGCTCGTATCGGAGACGACCTTGTCTTATCTGTGAAAGCG 7610  
QY 6962 CGGGAGCCAAAGAGACGAGCGAGCCCTACGGGCTTCACGAGGCTATGACTAGATACT 7021  
Db 7611 CGGGAGCCAAAGAGACGAGCGAGCCCTACGGGCTTCACGAGGCTATGACTAGATACT 7670  
QY 7022 CTGCCCCCTTGGGACCCGCCCAACACAGATACGACTTGGAGTTGATAATCATGCT 7081  
Db 7671 CTGCCCCCTTGGGACCCGCCCAACACAGATACGACTTGGAGTTGATAATCATGCT 7730  
QY 7082 CCTCAATGTCTCAGTCCGCGACGATGATCTGGCAAAAGGTTACTATCTCACCCGTG 7141  
Db 7731 CCTCAATGTCTCAGTCCGCGACGATGATCTGGCAAAAGGTTACTATCTCACCCGTG 7790  
QY 7142 ACCCCACACCCCTTGGCGGGGTGCGTGGGAGACAGCTAGACACACTCCAGTCAATT 7201  
Db 7791 ACCCCACACCCCTTGGCGGGGTGCGTGGGAGACAGCTAGACACACTCCAGTCAATT 7850  
QY 7202 CCTGCTAGGCAACATCATGATGATGCGCCACCTTGTGGGCAAGGATGATCCTGATGA 7261  
Db 7851 CCTGCTAGGCAACATCATGATGATGCGCCACCTTGTGGGCAAGGATGATCCTGATGA 7910

QY 7262 CTCATTTCTTCTCCATCTCTTAGCTCAGGACAACTTGAAAAGCCCTAGATTGTGACA 7321  
Db 7911 CTCATTTCTTCTCCATCTCTTAGCTCAGGACAACTTGAAAAGCCCTAGATTGTGACA 7970  
QY 7322 TCTACGGGCGCTGTTTACTCCATTGAGCCACTTGACTACCTCAGATCAATTCAACGACTCC 7381  
Db 7971 TCTACGGGCGCTGTTTACTCCATTGAGCCACTTGACTACCTCAGATCAATTCAACGACTCC 8030  
QY 7382 ATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTT 7441  
Db 8031 ACGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTT 8090  
QY 7442 CATGCTCAGAAACTTGGGGTACCGCTTGGAGTCTGGAGACATCGGGCCAGAAGTG 7501  
Db 8091 CATGCTCAGAAACTTGGGGTACCGCTTGGAGTCTGGAGACATCGGGCCAGAAGTG 8150  
QY 7502 TCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTCA 7561  
Db 8151 TCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTCA 8210  
QY 7562 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCNAATCCGGCTCGCTCCAGTTGGATT 7621  
Db 8211 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCNAATCCGGCTCGCTCCAGTTGGATT 8270  
QY 7622 TATCCAGCTGTGTTGCTGCTGTACAGCGGGGAGACATATATACAGCCTGTCTCGTG 7681  
Db 8271 TATCCAGCTGTGTTGCTGCTGTACAGCGGGGAGACATATATACAGCCTGTCTCGTG 8330  
QY 7682 CCGACCCCGCTGTTCAATGTGTGCTTACTCTTCTGTAGGGGTAGGCATCTATC 7741  
Db 8331 CCGACCCCGCTGTTCAATGTGTGCTTACTCTTCTGTAGGGGTAGGCATCTATC 8390  
QY 7742 TACTCCCCAACCGATGAACGGGACCTTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 7801  
Db 8391 TACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 8450  
QY 7802 TTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7861  
Db 8451 TTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8510  
QY 7862 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7921  
Db 8511 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8570  
QY 7922 AGCTGTGAAGCTCGGTGAGCGCTTGACTGAGAGAGTGTGATGACTGAGCTCTGCA 7981  
Db 8571 AGCTGTGAAGCTCGGTGAGCGCTTGACTGAGAGAGTGTGATGACTGAGCTCTGCA 8630  
QY 7982 GATCAAGT 7989  
Db 8631 GATCAAGT 8638

RESULT 10  
US-10-029-907-7  
; Sequence 7, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-7

Query Match      90.7%; Score 7248.2; DB 4; Length 8638;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 7960; Conservative 0; Mismatches 18; Indels 671; Gaps 3;

QY 1 GCGAGCCCGGATGGGGGGGACACTCCACATAGATCACTCCCTCTGTGAGAACTACTG 60
Db 1 GCGAGCCCGGATGGGGGGGACACTCCACATAGATCACTCCCTCTGTGAGAACTACTG 60
QY 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGCNGAGCCCTCCAGGAC 120
Db 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGCNGAGCCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG 180
QY 181 GAGGACCGGCTCTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240
Db 181 GAGGACCGGCTCTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240
QY 241 GCGAGACTGCTAGCGAGTGTCTGGGTGCGGAAAGGCTTGTGCTACTGCTGATAGG 300
Db 241 GCGAGACTGCTAGCGAGTGTCTGGGTGCGGAAAGGCTTGTGCTACTGCTGATAGG 300
QY 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTGTAAGCCGTGCACCATGAGCAGCAATTCCTAAAC 360
Db 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTGTAAGCCGTGCACCATGAGCAGCAATTCCTAAAC 360
QY 361 CTCAGAGAAACCAAGAGCGCGCCATGATGAACAGATGATGACAGCGAGTTCTC 420
Db 361 CTCAGAGAAACCAAGAGCGCGCCATGATGAACAGATGATGACAGCGAGTTCTC 420
QY 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACAAAGCAATCGGCTGT 480
Db 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACAAAGCAATCGGCTGT 480
QY 481 CTGATGCCCGCTGTTCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540
Db 481 CTGATGCCCGCTGTTCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540
QY 541 ACTGTCCGGTCCCTGATGAATCACTGAGAGCGAGGCGGCTATCGTGGCTGGCCA 600
Db 541 ACTGTCCGGTCCCTGATGAATCACTGAGAGCGAGGCGGCTATCGTGGCTGGCCA 600
QY 601 CGAGGGCGCTTCTTCGCGAGCTGTCTGACGTTGTCTCACTGAAGCGGAAGGACTGGC 660
Db 601 CGAGGGCGCTTCTTCGCGAGCTGTCTGACGTTGTCTCACTGAAGCGGAAGGACTGGC 660
QY 661 TGTATTGGGCGAAGTGCAGGGGAGGATCTCTGTCACTCACTTGTCTTCCGCGAGA 720
Db 661 TGTATTGGGCGAAGTGCAGGGGAGGATCTCTGTCACTCACTTGTCTTCCGCGAGA 720
QY 721 AGATATCCATATGGCTGATGCAATCGGCGGCTGATACGCTTGTATCGGCTTACCTGCC 780
Db 721 AGATATCCATATGGCTGATGCAATCGGCGGCTGATACGCTTGTATCGGCTTACCTGCC 780
QY 781 CATTCGACCAACAGCGAAACATCGCATCGAGGAGCAGTACTCGATGGAAGCGGTC 840
Db 781 CATTCGACCAACAGCGAAACATCGCATCGAGGAGCAGTACTCGATGGAAGCGGTC 840
QY 841 TTGTCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCGAACTGTTCG 900
Db 841 TTGTCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCGAACTGTTCG 900
QY 901 CCAGGCTCAAGCGCGCATGCCGACGCGGAGGATCTCGTGTGA CCGATGGCGATGCCT 960
Db 901 CCAGGCTCAAGCGCGCATGCCGACGCGGAGGATCTCGTGTGA CCGATGGCGATGCCT 960
QY 961 GCTTGGCGAATATCATGTTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGC 1020

```

```

Db 961 GCTTGGCGAATATCATGTTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGC 1020
QY 1021 TGGGTGTGGCGAACCGCTATCAGGACATAGCGTGTGGCTACCGCTGATATGCTGAAGAGC 1080
Db 1021 TGGGTGTGGCGAACCGCTATCAGGACATAGCGTGTGGCTACCGCTGATATGCTGAAGAGC 1080
QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCCG 1140
Db 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCCG 1140
QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTT-----TAAA 1188
Db 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTTCGCGGCCAGATGTAA 1200
QY 1189 CAGACCAACAACGGTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCCCT 1248
Db 1201 CAGACCAACAACGGTTCCTCTAGCGGGATCAATTCG-----CCCCCCCCCT 1249
QY 1249 AAGCTTACTGGCGAAGCGCTTGGATTAAGCGCGGTGCGTGTGTCTATATGTTATTT 1308
Db 1250 AAGCTTACTGGCGAAGCGCTTGGATTAAGCGCGGTGCGTGTGTCTATATGTTATTT 1309
QY 1309 TCCACCATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTG 1368
Db 1310 TCCACCATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTG 1369
QY 1369 ACAGCATTCCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGGAATGTC 1428
Db 1370 ACAGCATTCCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGGAATGTC 1429
QY 1429 GTGAGGAGCAGTTCTCTGGAAGTCTTTCGAAGCAACACGCTCTGTAGCGACCCCT 1488
Db 1430 GTGAGGAGCAGTTCTCTGGAAGTCTTTCGAAGCAACACGCTCTGTAGCGACCCCT 1489
QY 1489 TGCAAGCAGCGGAAACCCCGACCTGCGGACAGGTGCTCTGCGGCCAAAGCCACGTGTA 1548
Db 1490 TGCAAGCAGCGGAAACCCCGACCTGCGGACAGGTGCTCTGCGGCCAAAGCCACGTGTA 1549
QY 1549 TAAGTATACCTGTCAAGAGCGGCAACCCCGACCTGCGGACAGGTGCTCTGCGGCCAAAG 1608
Db 1550 TAAGTATACCTGTCAAGAGCGGCAACCCCGACCTGCGGACAGGTGCTCTGCGGCCAAAG 1609
QY 1609 GAAAGTCAATAGCTCTCTCAAGCTATTCAACAAAGGCTGAAGATGCCCAAG 1668
Db 1610 GAAAGTCAATAGCTCTCTCAAGCTATTCAACAAAGGCTGAAGATGCCCAAG 1669
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCTCTCGGTGCACATGCTTTACATGTTTAG 1728
Db 1670 GTACCCCATTTGATGGATCTGATCTGGGCTCTCGGTGCACATGCTTTACATGTTTAG 1729
QY 1729 TCAGGTTAAAAAAGCTCTAGCCCCCGGAAACCAAGGAGCGTGTCTTCTTTGAAAAA 1788
Db 1730 TCAGGTTAAAAAAGCTCTAGCCCCCGGAAACCAAGGAGCGTGTCTTCTTTGAAAAA 1789
QY 1789 CAGGATAATACCATG----- 1803
Db 1790 CAGGATAATACCATG----- 1803
QY 1804 ----- 1803
Db 1850 CTGATACTCTTACCTTGTACCGCACTATAAGTGTCTCTCGTAGGCTCATATGTTG 1909
QY 1804 ----- 1803
Db 1910 TTACAATATTTTATCACCAGGCGCGAGGCACACTTGTCAAGTGTGGATCCCCCCCCCAAC 1969
QY 1804 ----- 1803
Db 1970 GTTCGGGGGGCGCGATGCCGTCTCTCTCAGTGGCGGATCCACCCAGAGCTAATC 2029
QY 1804 ----- 1803

```

Db 2030 TTTACCATCAGCAAAATCTTGCTGCCATACCTCGGTCCACTCATGGTGTCCAGGCTGGT 2089  
QY 1804 ----- 1803  
Db 2090 ATAACCAAGTGCCGTACTTGCTGCGGCACACGGGCTCATTCGTGCATGATGTGGTG 2149  
QY 1804 ----- 1803  
Db 2150 CGGAAGTTGCTGGGGTCAATATGTCACAAATGGCTCTCATGAAGTTGGCCGCACTGACA 2209  
QY 1804 ----- 1803  
Db 2210 GGTACGTACTGTTATGAACATCTCACCCCACTGCGGGAGCTGGGCCACACGCGGGCTACGA 2269  
QY 1804 ----- 1803  
Db 2270 GACCTTGGGTGGCAGTTGAGCCGCTGCTTCTCTGATATGGACCAAGGTTATCACC 2329  
QY 1804 ----- 1803  
Db 2330 TGGGGGCGACACCGCGGCTGTGGGACATCATCTTGGGCTGCCCGTCTCCGCCGCG 2389  
QY 1804 ----- 1803  
Db 2390 AGGGGAGGGAGATACATCTGGGACCGGCAGACAGCCTTGAAGGCGAGGGGTGGCGACTC 2449  
QY 1804 ---GCGCCTATTACGGCTACTCCCAACAGAGCGAGGCTACTTGGCTGCATCATCT 1860  
Db 2450 CTCGGCCCTATTACGGCCCTACTCCCAACAGAGCGAGGCTACTTGGCTGCATCATCTACC 2509  
QY 1861 AGCCTCACAGGCGGGGACAGGAACAGAGTCCAGGGGAGGCTCCAAAGTGTCTCCACCGCA 1920  
Db 2510 AGCCTCACAGGCGGGGACAGGAACAGAGTCCAGGGGAGGCTCCAAAGTGTCTCCACCGCA 2569  
QY 1921 ACACAACTTCTTGGGACCTGCTGCTAAATGGCGTGTGGAGCTGTCTATCATGGTGCC 1980  
Db 2570 ACACAACTTCTTGGGACCTGCTGCTAAATGGCGTGTGGAGCTGTCTATCATGGTGCC 2629  
QY 1981 GGCTCAAGAGCCCTTGGCGGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 2630 GGCTCAAGAGCCCTTGGCGGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2689  
QY 2041 CAGGACCTCGTGGGTGCAAGCGCCCGCGGGCGCTTCCTTGACACCACTGACACCTGC 2100  
Db 2690 CAGGACCTCGTGGGTGCAAGCGCCCGCGGGCGCTTCCTTGACACCACTGACACCTGC 2749  
QY 2101 GCGAGCTCGAGCTTTACTTGGTCAAGAGGATGCCGATGTGATCCGGTGCGCGCGG 2160  
Db 2750 GCGAGCTCGAGCTTTACTTGGTCAAGAGGATGCCGATGTGATCCGGTGCGCGCGG 2809  
QY 2161 GCGGACAGCGGGGAGCTACTCTCCCGCAGGCCCGCTCTCTACTTGAAGGCTCTTCG 2220  
Db 2810 GCGGACAGCGGGGAGCTACTCTCCCGCAGGCCCGCTCTCTACTTGAAGGCTCTTCG 2869  
QY 2221 GCGGTCCACTGCTCTGCGCCCTCGGGGCACTGTGGGCACTTTTCGGGCTGCCGTGTGC 2280  
Db 2870 GCGGTCCACTGCTCTGCGCCCTCGGGGCACTGTGGGCACTTTTCGGGCTGCCGTGTGC 2929  
QY 2281 ACCGAGGGGTGGAGGGGTGACCTTTGTATCCCGTGCAGTCTATGGAAACCACTATG 2340  
Db 2930 ACCGAGGGGTGGAGGGGTGACCTTTGTATCCCGTGCAGTCTATGGAAACCACTATG 2989  
QY 2341 CGGTCCCGGCTTTCAGCGACAACTCGTCCCTCGGCGGTACCGCACATTCACAGTG 2400  
Db 2990 CGGTCCCGGCTTTCAGCGACAACTCGTCCCTCGGCGGTACCGCACATTCACAGTG 3049  
QY 2401 GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCCGGCTGCCGTATGCA 2460  
Db 3050 GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCCGGCTGCCGTATGCA 3109  
QY 2461 GCCAAGGTATAGGTGCTTGTCTGAAACCGGCTCGCGCCACCTAGGTTCGGG 2520  
Db 3110 GCCAAGGTATAGGTGCTTGTCTGAAACCGGCTCGCGCCACCTAGGTTCGGG 3169

QY 2521 GCGTATATGTCTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTGAAGCACTC 2580  
Db 3170 GCGTATATGTCTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGTGAAGCACTC 3229  
QY 2581 ACCAGGGTGGCCCGCATCACGTACTCCACCTATGGCAAGTTTCTTTCGCGAGCGTGTGC 2640  
Db 3230 ACCAGGGTGGCCCGCATCACGTACTCCACCTATGGCAAGTTTCTTTCGCGAGCGTGTGC 3289  
QY 2641 TCTGGGGCGGCTATGACATCATATATGTATGATGATGCCACTAATGACTCGACCACT 2700  
Db 3290 TCTGGGGCGGCTATGACATCATATATGTATGATGATGCCACTAATGACTCGACCACT 3349  
QY 2701 ATCTTGGGCATCGGCACAGTCTCGACCAAGCGAGCGGTGGAGCGGCTCTCTGTG 2760  
Db 3350 ATCTTGGGCATCGGCACAGTCTCGACCAAGCGAGCGGTGGAGCGGCTCTCTGTG 3409  
QY 2761 CTCGCCACCGCTACGCTTCCGGGATCGGTACCGTCCACATCCAAACATCGAGGAGTG 2820  
Db 3410 CTCGCCACCGCTACGCTTCCGGGATCGGTACCGTCCACATCCAAACATCGAGGAGTG 3469  
QY 2821 GCTCTGTCAGGACTGGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGCACTC 2880  
Db 3470 GCTCTGTCAGGACTGGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGCACTC 3529  
QY 2881 AAGGGGGGAGGCACCTCATTTTCTGCCATTCRAAGAGAAATGTATGATGCTCCCGG 2940  
Db 3530 AAGGGGGGAGGCACCTCATTTTCTGCCATTCRAAGAGAAATGTATGATGCTCCCGG 3589  
QY 2941 AAGTGTTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCGTC 3000  
Db 3590 AAGTGTTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCGTC 3649  
QY 3001 ATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
Db 3650 ATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATGACGGGCTTACC 3709  
QY 3061 GCGGATTCGACTCAGTGATGACTGCAATACATGTGTCAACCAGACAGTGCATTCAGC 3120  
Db 3710 GCGGATTCGACTCAGTGATGACTGCAATACATGTGTCAACCAGACAGTGCATTCAGC 3769  
QY 3121 CTGACCCGACCTTTCACANTGAGACGACCGCTGCCACAGAGCGGGTGTACGCTCG 3180  
Db 3770 CTGACCCGACCTTTCACANTGAGACGACCGCTGCCACAGAGCGGGTGTACGCTCG 3829  
QY 3181 CAGCGGCGAGCGAGGACTGGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGGA 3240  
Db 3830 CAGCGGCGAGCGAGGACTGGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGGA 3889  
QY 3241 GAACGGCCCTCGGCATGTTCTCGATTCTCTCGGTTCTGTGCGAGTGCTATGACGGGCTGT 3300  
Db 3890 GAACGGCCCTCGGCATGTTCTCGATTCTCTCGGTTCTGTGCGAGTGCTATGACGGGCTGT 3949  
QY 3301 GCTTGTACGAGCTCAGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACTTAAACACA 3360  
Db 3950 GCTTGTACGAGCTCAGCCCGCGGAGACCTCAGTTAGTTTGGGGCTTACTTAAACACA 4009  
QY 3361 CAGGGTTCGCCGTCTGCCAGGACCACTGAGTTCTGGGAGAGGCTTCTTACAGGCTC 3420  
Db 4010 CAGGGTTCGCCGTCTGCCAGGACCACTGAGTTCTGGGAGGGGCTTCTTACAGGCTC 4069  
QY 3421 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGAGCAACTTCCCTTAC 3480  
Db 4070 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGAGCAACTTCCCTTAC 4129  
QY 3481 CTGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGTTCCACCTCCATCGTGGAC 3540  
Db 4130 CTGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGTTCCACCTCCATCGTGGAC 4189  
QY 3541 CAATGTGGAAGTGTCTCATACGCTAAGCCTACGCTGACGGGCCAACGCCCTGTG 3600  
Db 4190 CAATGTGGAAGTGTCTCATACGCTAAGCCTACGCTGACGGGCCAACGCCCTGTG 4249

QY 3601 TATAGCTGGAGCGGTTCAABAACGAGGTTACTACACACACCCCAATACCAATACATC 3660  
DB |||||  
QY 4250 TATAGCTGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAATACCAATACATC 4309  
DB |||||  
QY 3661 ATGGCATGATCGGCTGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGA 3720  
DB |||||  
QY 4310 ATGGCATGATCGGCTGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGA 4369  
DB |||||  
QY 3721 GTCTAGAGCTCTGGCCGCGTATTCCTTGACAAACAGGAGCGGTGCTATTTGGGACGG 3780  
DB |||||  
QY 4370 GTCTAGAGCTCTGGCTGCTATTTGCTGACAAACAGGAGCGGTGCTATTTGGGACGG 4429  
DB |||||  
QY 3781 ATCATCTCTCGGAAACCGGCATATTCCGACAGAGGAGTCTTTACCGGGAGTTC 3840  
DB |||||  
QY 4430 ATCACTCTGCGGAAGCGGCATATTCCGACAGGAAAGTCTTTACCGGGAGTTC 4489  
DB |||||  
QY 3841 GATGAGTGAAGAGTGCCTCTACACCTCCCTTACATCGAAACAGGGAATGCAGCTCGCC 3900  
DB |||||  
QY 4490 GATGAGTGAAGAGTGCCTCTACACCTCCCTTACATCGAAACAGGGAATGCAGCTCGCC 4549  
DB |||||  
QY 3901 GAACAAATTCAAACAGAGCAATCGGTTGTCGAAACAGCCACCAAGCAAGCGAGGCT 3960  
DB |||||  
QY 4550 GAACAAATTCAAACAGAGCAATCGGTTGTCGAAACAGCCACCAAGCAAGCGAGGCT 4609  
DB |||||  
QY 3961 GCTGCTCCGCTGCGTAATCCAGTGGCGAGCCCTCGAGCCCTCTGGGCGAGCATATG 4020  
DB |||||  
QY 4610 GCTGCTCCGCTGCGTAATCCAGTGGCGAGCCCTCGAGCCCTCTGGGCGAGCATATG 4669  
DB |||||  
QY 4021 TGGAAATTCATCAGCGGGATCAAAATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
DB |||||  
QY 4670 TGGAAATTCATCAGCGGGATCAAAATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCCC 4729  
DB |||||  
QY 4081 GCGATAGCATCACTGATGGAATTCACAGCTCTATCACAGCCCGCTCACCAACCAAT 4140  
DB |||||  
QY 4730 GCGATAGCATCACTGATGGAATTCACAGCTCTATCACAGCCCGCTCACCAACCAAT 4789  
DB |||||  
QY 4141 ACCCTCTGTTTAAATCTCTGGGGATGGGTGGCGCCCACTGCTCTCCCGAGCT 4200  
DB |||||  
QY 4790 ACCCTCTGTTTAAATCTCTGGGGATGGGTGGCGCCCACTGCTCTCCCGAGCT 4849  
DB |||||  
QY 4201 GCTTCTGCTTTTCTAGCGCCGCTATCGCTGAGCGGCTGTTGGCAGCATAGGCCCTGGG 4260  
DB |||||  
QY 4850 GCTTCTGCTTTTCTAGCGCCGCTATCGCTGAGCGGCTGTTGGCAGCATAGGCCCTGGG 4309  
DB |||||  
QY 4261 AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGCGAGCGGCTCGTGGCC 4320  
DB |||||  
QY 4910 AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGCGAGCGGCTCGTGGCC 4369  
DB |||||  
QY 4321 TTTAAGGTATGAGCGCGAGATGCCCTCCACCGAGGACCTGTTAACTCTCCCTGCT 4380  
DB |||||  
QY 4970 TTTAAGGTATGAGCGCGAGATGCCCTCCACCGAGGACCTGTTAACTCTCCCTGCT 5029  
DB |||||  
QY 4381 ATCTCTCTCCCTGGCCCTAGTCTGCGGGTCTGTCGCGCAGCATATCTGGCTCGGCAC 4440  
DB |||||  
QY 5030 ATCTCTCTCCCTGGCCCTAGTCTGCGGGTCTGTCGCGCAGCATATCTGGCTCGGCAC 5089  
DB |||||  
QY 4441 GTGGCCCGAGGGAGGGGCTGTCAGTGAATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
DB |||||  
QY 5090 GTGGCCCGAGGGAGGGGCTGTCAGTGAATGAACCGGCTGATAGCGTTGCTTCGCGG 5149  
DB |||||  
QY 4501 GGTAAACAGCTCTCCCGCAGCATATGTCCTGAGAGCGGCTCGAGCAGTGTCACT 4560  
DB |||||  
QY 5150 GGTAAACAGCTCTCCCGCAGCATATGTCCTGAGAGCGGCTCGAGCAGTGTCACT 5209  
DB |||||  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTCTGAAAGGCTTCACCAAGTGGATCAAC 4620  
DB |||||  
QY 5210 CAGATCTCTCTAGTCTTACCATCACTCAGCTCTGAAAGGCTTCACCAAGTGGATCAAC 5269  
DB |||||  
QY 4621 GAGGACTGTCCACGCCATGCTCCGCTCGTGGCTAAGAGATGTTGGGATTTGGATATGC 4680  
DB |||||  
QY 5270 GAGGACTGTCCACGCCATGCTCCGCTCGTGGCTAAGAGATGTTGGGATTTGGATATGC 5329  
DB |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATTTGCCGGA 4740

DB |||||  
QY 5330 ACGGTGTTGACTGATTTCAAGGCTTGGCTCCAGTCCAAAGCTCTGCGCGGATTTGCCGGA 5389  
DB |||||  
QY 4741 GTCCCTCTTCTTCATGTCAAGTGGGTACAAGGGAGTCTGGCGGGCGACGCAATCATG 4800  
DB |||||  
QY 5390 GTCCCTCTTCTTCATGTCAAGTGGGTACAAGGGAGTCTGGCGGGCGACGCAATCATG 5449  
DB |||||  
QY 4801 CAAACACCTGCCCCATGTGGAGCAGATCAACGGACATGTGAANAACGGTTCCATGAG 4860  
DB |||||  
QY 5450 CAAACACCTGCCCCATGTGGAGCAGATCAACGGACATGTGAANAACGGTTCCATGAG 5509  
DB |||||  
QY 4861 ATCTGTTGGGCTTAGACCTGTAGTAACACGTGGCAATTTCCCAATTAACGGGTAC 4920  
DB |||||  
QY 5510 ATCTGTTGGGCTTAGACCTGTAGTAACACGTGGCAATTTCCCAATTAACGGGTAC 5569  
DB |||||  
QY 4921 ACCACGGGCCCCCTGACGCCCCCTCCCGGCGCAAAATTTATTTAGGGCGTGTGGCGGTG 4980  
DB |||||  
QY 5570 ACCACGGGCCCCCTGACGCCCCCTCCCGGCGCAAAATTTATTTAGGGCGTGTGGCGGTG 5629  
DB |||||  
QY 4981 GCTGCTGAGGAGTACGTGGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
DB |||||  
QY 5630 GCTGCTGAGGAGTACGTGGAGGTTACGCGAGTGGGGATTTCCACTACGTGACGGGCATG 5689  
DB |||||  
QY 5041 ACCACTGACAAAGTAAAGTCCCGTGTAGTTCCGGCCCCCGAAATTTCTTTCAGAAAGTG 5100  
DB |||||  
QY 5690 ACCACTGACAAAGTAAAGTCCCGTGTAGTTCCGGCCCCCGAAATTTCTTTCAGAAAGTG 5749  
DB |||||  
QY 5101 GATGGGTGCGGTTGACAGGTACGCTCCAGCGTGCAAAACCCCTCTACGGGAGGAGTTC 5160  
DB |||||  
QY 5750 GATGGGTGCGGTTGACAGGTACGCTCCAGCGTGCAAAACCCCTCTACGGGAGGAGTTC 5809  
DB |||||  
QY 5161 AATTTCTGCTGCTGAGTCAATCAATACCTGTTGGGTCAAGCTCCCATGCGAGCCGGA 5220  
DB |||||  
QY 5810 AATTTCTGCTGCTGAGTCAATCAATACCTGTTGGGTCAAGCTCCCATGCGAGCCGGA 5869  
DB |||||  
QY 5221 CCGGAGTACAGTGTCTACTTCCATGTCAACGACCCCTCCACATTAACGGCGGAGACG 5280  
DB |||||  
QY 5870 CTGAGCTGAGAGTGTCTACTTCCATGTCAACGACCCCTCCACATTAACGGCGGAGACG 5929  
DB |||||  
QY 5281 GCTAAGCTGAGCTGGCAGGGGATCTCCCCCTCTCTTGGCCAGCTCATCAGTAGCCAG 5340  
DB |||||  
QY 5930 GCTAAGCTGAGCTGGCAGGGGATCTCCCCCTCTCTTGGCCAGCTCATCAGTAGCCAG 5989  
DB |||||  
QY 5341 CTGCTGCGCCTTCTTGAAGGCAATGCACTACCCGTCAATGATCCCGCGAGCGTGAC 5400  
DB |||||  
QY 5990 CTGCTGCGCCTTCTTGAAGGCAATGCACTACCCGTCAATGATCCCGCGAGCGTGAC 6049  
DB |||||  
QY 5401 CTCTATCGAGGCCAACTCTCTGTGGCGCAGGAGATGGCGGGAAATCAACCCCGTGGAG 5460  
DB |||||  
QY 6050 CTCTATCGAGGCCAACTCTCTGTGGCGCAGGAGATGGCGGGAAATCAACCCCGTGGAG 6109  
DB |||||  
QY 5461 TCAGAAATTAAGTATGTAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGTATG 5520  
DB |||||  
QY 6110 TCAGAAATTAAGTATGTAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGATGAG 6169  
DB |||||  
QY 5521 AGGAAAGTATCGTTCCGCGGAGATCTCGGAGGTTCAGAAATTTCCCTCGAGCGATG 5580  
DB |||||  
QY 6170 AGGAAAGTATCGTTCCGCGGAGATCTCGGAGGTTCAGGAAATTTCCCTCGAGCGATG 6229  
DB |||||  
QY 5581 CCCATATGGGCACGCCCGGATTTACACCCCTCCACTGTTAGAGTTCCTGGAAGGACCCCGGAC 5640  
DB |||||  
QY 6230 CCCATATGGGCACGCCCGGATTTACACCCCTCCACTGTTAGAGTTCCTGGAAGGACCCCGGAC 6289  
DB |||||  
QY 5641 TAGCTCCCTTCCAGTGTGATACACGGGTGTCCATTTGCGGCTTCGCAAGGCCCTCCGATATCCA 5700  
DB |||||  
QY 6290 TAGCTCCCTTCCAGTGTGATACACGGGTGTCCATTTGCGGCTTCGCAAGGCCCTCCGATATCCA 6349  
DB |||||  
QY 5701 CTTCCACGGAAGGAGGACGGTGTCTGTGATCAATCTACCGTGTCTTCTGCTTGGCG 5760  
DB |||||  
QY 6350 CTTCCACGGAAGGAGGACGGTGTCTGTGATCAATCTACCGTGTCTTCTGCTTGGCG 6409  
DB |||||  
QY 5761 GAGCTCGCCACAAAGACCTTTCGCGAGCTCCGAATGCTCGGCGCTCGACACGGCAGCGCA 5820  
DB |||||





```
QY 7981 AGATCAAGT 7989
Db 8630 AGATCAAGT 8638

RESULT 11
US-10-029-907-25
; Sequence 25, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-25

Query Match 90.7%; Score 7247.2; DB 4; Length 8638;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 7959; Conservative 0; Mismatches 18; Indels 671; Gaps 3;

QY 2 CCAGCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAATCTACTGT 61
Db 2 CCAGCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAATCTACTGT 61

QY 62 CTTACGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTGCTGAGCGCTCCAGGACC 121
Db 62 CTTACGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTGCTGAGCGCTCCAGGACC 121

QY 122 CCCCTCCGGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATGCGCAGG 181
Db 122 CCCCTCCGGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATGCGCAGG 181

QY 182 ACGACCGGTCTCTTCTTGGATCAACCCGCTCAATGCCTGGAGATTGGCGCTGCCCGG 241
Db 182 ACGACCGGTCTCTTCTTGGATCAACCCGCTCAATGCCTGGAGATTGGCGCTGCCCGG 241

QY 242 CGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAGGCTTGTGTACTGTGCTGTAGGG 301
Db 242 CGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAGGCTTGTGTACTGTGCTGTAGGG 301

QY 302 TGTTCGAGTGTCCCGGAGGTCTGTAGACCGTGCACATGAGCAGCAATCTTAACC 361
Db 302 TGTTCGAGTGTCCCGGAGGTCTGTAGACCGTGCACATGAGCAGCAATCTTAACC 361

QY 362 TCAAGAAAAACCAAGGGCGGCCATGATTGAACAAGATGATTGACGCGAGTTCTCC 421
Db 362 TCAAGAAAAACCAAGGGCGGCCATGATTGAACAAGATGATTGACGCGAGTTCTCC 421

QY 422 GCGCGTGGGTGGAGAGGTATTCCGCTATGATGCTGGGCAACAGAGATCGGCTGCTC 481
Db 422 GCGCGTGGGTGGAGAGGTATTCCGCTATGATGCTGGGCAACAGAGATCGGCTGCTC 481

QY 482 TGATGCGCGCTTCTCCGCTGTGACGCGAGGGCGCGGTTCTTTTGTCAAGACCGA 541
Db 482 TGATGCGCGCTTCTCCGCTGTGACGCGAGGGCGCGGTTCTTTTGTCAAGACCGA 541

QY 542 CTTGTCCGCTGCCCTGAATGAATCTGAGGACGAGGCGCGGCTATCGTGGCTGGCCAC 601
Db 542 CTTGTCCGCTGCCCTGAATGAATCTGAGGACGAGGCGCGGCTATCGTGGCTGGCCAC 601

602 GACGGGCGTTCTTGGCAGCTGTCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCT 661
Db 602 GACGGGCGTTCTTGGCAGCTGTCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCT 661

662 GCTATTGGCGGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCCAGAA 721
Db 662 GCTATTGGCGGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCCAGAA 721

722 AGTATCCATCATGGCTGATGAAATCGCGGCTGTGATACGTTGATCCGGCTACCTGCC 781
Db 722 AGTATCCATCATGGCTGATGAAATCGCGGCTGTGATACGTTGATCCGGCTACCTGCC 781

782 ATTGACACCAAGCGAATCGATCGAGGACGATCTCGATCGAGGACGATCTCGATCGAGGACGATCT 841
Db 782 ATTGACACCAAGCGAATCGATCGAGGACGATCTCGATCGAGGACGATCTCGATCGAGGACGATCT 841

842 TGTTCGATCAGGATGATCTTGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCG 901
Db 842 TGTTCGATCAGGATGATCTTGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCG 901

902 CAGGCTCAAGGCGGATGCCCGGAGGATCTCGTGTGATGCCATCGGATGCGCTG 961
Db 902 CAGGCTCAAGGCGGATGCCCGGAGGATCTCGTGTGATGCCATCGGATGCGCTG 961

962 CTTGCGGATATCATGGTGGAAAAATGGCGCTTTCTGGATTCAATGATGTCGCGGCT 1021
Db 962 CTTGCGGATATCATGGTGGAAAAATGGCGCTTTCTGGATTCAATGATGTCGCGGCT 1021

1022 GGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATGCTGAAGAGCT 1081
Db 1022 GGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATGCTGAAGAGCT 1081

1082 TGGCGCGAATGGCTGACCGCTTCTCTGTGTTCAGGTATCGCGCTCCGATTCGA 1141
Db 1082 TGGCGCGAATGGCTGACCGCTTCTCTGTGTTCAGGTATCGCGCTCCGATTCGA 1141

1142 GCGCATGCGCTTCTATCGCGCTTCTTGACGAGTCTTCTGAGTTGCGGCCAGATGTAAC 1189
Db 1142 GCGCATGCGCTTCTATCGCGCTTCTTGACGAGTCTTCTGAGTTGCGGCCAGATGTAAC 1189

1190 AGACACAAACGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTA 1249
Db 1202 AGACACAAACGTTTCCCTCTAGCGGATCAATTCGCGGCCCTCTCCCTCCCCCCCCCTA 1250

1250 AGCTTACTGCGGCGAAGCGCTTGGAAATAGGCGCGGTGCGTTTGTCTATATGTTATTT 1309
Db 1251 AGCTTACTGCGGCGAAGCGCTTGGAAATAGGCGCGGTGCGTTTGTCTATATGTTATTT 1310

1310 CCACCATATTGCGGCTTTTGGCAATGTAGGCGCGGAAACCTGCGCCCTGCTTCTTGA 1369
Db 1311 CCACCATATTGCGGCTTTTGGCAATGTAGGCGCGGAAACCTGCGCCCTGCTTCTTGA 1370

1370 CGAGCATTTCTAGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1429
Db 1371 CGAGCATTTCTAGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1430

1430 TGAAGAGACGATTCCTCTGGAAGCTTTTGAAGACAAACAACTGTGTAGGACCCCTTT 1489
Db 1431 TGAAGAGACGATTCCTCTGGAAGCTTTTGAAGACAAACAACTGTGTAGGACCCCTTT 1490

1490 GCAGGCGGGAACCCCCCAGCTGCGGACAGGTGCTCTGCGGCCAAAGACCGCTGAT 1549
Db 1491 GCAGGCGGGAACCCCCCAGCTGCGGACAGGTGCTCTGCGGCCAAAGACCGCTGAT 1550

1550 AAGATACACCTCAAGGCGGCAACCCAGTGCACCTTGTGAGTTGATGTTGCTG 1609
Db 1551 AAGATACACCTCAAGGCGGCAACCCAGTGCACCTTGTGAGTTGATGTTGCTG 1610

1610 AAAGAGTCAAAATGGCTCTCTCAAGCTATTCAACAAAGGGGCTGAAGATGCCCAAGG 1669
Db 1611 AAAGAGTCAAAATGGCTCTCTCAAGCTATTCAACAAAGGGGCTGAAGATGCCCAAGG 1670

1670 TACCCCAATTGATGGAATCTGATCTGGGGCTCGGTGCAATGCTTTTATCATGTTGTTAGT 1729
```

Db 1671 TACCCATTGATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTAGT 1730  
QY 1730 CGAGGTTAAAAACGTCCTAGGCCCCCGAAACACGAGGACGTGGTTTCCCTTTGAAAAAC 1789  
Db 1731 CGAGGTTAAAAACGTCCTAGGCCCCCGAAACACGAGGACGTGGTTTCCCTTTGAAAAAC 1790  
QY 1790 ACGATAATACCATG----- 1803  
Db 1791 ACGATAATACCATGACCGGAGATGGCAGCATCGTGGAGGCGGGTTTTCGTAGGTC 1850  
QY 1804 ----- 1803  
Db 1851 TGATACTCTTGACCTTGTACCGCACTATAAGCTGTTCCCTCGTAGGCTCATATGTTGGT 1910  
QY 1804 ----- 1803  
Db 1911 TACAATAATTTATACACGAGGCGGAGGACACTTGCAGAGTGTGGATCCCCCCTCAACG 1970  
QY 1804 ----- 1803  
Db 1971 TTCGGGGGCGCGATGCGGCTATCCTCCTCACGTGCGGATCCACCCAGAGCTAATCT 2030  
QY 1804 ----- 1803  
Db 2031 TTACCATCACCAAAATCTTGTCTGCCATACTCGGTCCACTCATGTGCTCCAGGCTGGTA 2090  
QY 1804 ----- 1803  
Db 2091 TAACAAAAGTGCCTACTTCTGTGCGCGCACACAGGCTCATTCGTGTCATGCTGTGTC 2150  
QY 1804 ----- 1803  
Db 2151 GGAAGTTGCTGGGGTCATTAATGTCAAATGGCTCTCATGAAGTTGGCGGCACTGACAG 2210  
QY 1804 ----- 1803  
Db 2211 GTACGTACGTTTATGACCATCTCACCCACTGCGGACTGGGCCACCGGGCTTACGAG 2270  
QY 1804 ----- 1803  
Db 2271 ACCTTGGGTGGCAGTTGAGCCCGTCGTCTTCTGTATATGAGACCAAGGTTATACCT 2330  
QY 1804 ----- 1803  
Db 2331 GGGGGGAGACACCGCGGCTGTGGGACATCATCTTGGGCTGCCCTCTCCGCCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGAGGGAGATACATCTGGACCGGCAGACAGCCTTTGAAGGGCAGGGTGGGACTCC 2450  
QY 1804 --GGCCTATTACGGCTACTCCCAACAGACGGAGGCTACTTGGCTGCACTACATA 1861  
Db 2451 TCGGGCTATTACGGCTACTCCCAACAGACGGAGGCTACTTGGCTGCACTACATA 2510  
QY 1862 GCCTCACAGGCGGACAGGAACAGGTCGAGGGGAGTCCAGTGTCTCCACCGCAA 1921  
Db 2511 GCCTCACAGGCGGACAGGAACAGGTCGAGGGGAGTCCAGTGTCTCCACCGCAA 2570  
QY 1922 CACAATCTTCTGTGCGCACTCGCTCAATGCGTGTGTGGACTGTCTATCATGTTGCCG 1981  
Db 2571 CACAATCTTCTGTGCGCACTCGCTCAATGCGTGTGTGGACTGTCTATCATGTTGCCG 2630  
QY 1982 GCTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC 2041  
Db 2631 GCTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC 2690  
QY 2042 AGGACTCGTGGGTGGAAGCGCCCGGGGCGGCTTCTTGACACCATGCACTGCG 2101  
Db 2691 AGGACTCGTGGGTGGAAGCGCCCGGGGCGGCTTCTTGACACCATGCACTGCG 2750  
QY 2102 CGAGCTCGGACCTTTACTTGTGTACAGGCAATGCCATGTCAATTCGGTGGCGCGGG 2161

Db 2751 CGAGCTCGGACCTTTACTTGTGTACGAAGCATGCGATGTCAATTCGGTGGCGCGGG 2810  
QY 2162 CGGACAGAGGGGAGACCTACTCTCCCCAGGCGGCTCTCTACTTTGAAGGCTCTTCGG 2221  
Db 2811 CGGACAGAGGGGAGACCTACTCTCCCCCGGCGGCTCTCTACTTTGAAGGCTCTTCGG 2870  
QY 2222 CGGTGCCACTCTCTGCCCTCGGGCACGCTGTGGGCACTTTTCGGGCTGCCGTGTGCA 2281  
Db 2871 CGGTGCCACTCTCTGCCCTCGGGCACGCTGTGGGCACTTTTCGGGCTGCCGTGTGCA 2930  
QY 2282 CCCGAGGGGTTGCGAAGCGGTGCACTTTGTACCGGTGAGTCTATGAAACCACTATGC 2341  
Db 2931 CCCGAGGGTTGCGAAGCGGTGCACTTTGTACCGGTGAGTCTATGAAACCACTATGC 2990  
QY 2342 GGTCCCCGGTCTTACGAGCAACTCGTCCCTCGGGCGTACCGGAGACATTCAGAGTGG 2401  
Db 2991 GGTCCCCGGTCTTACGAGCAACTCGTCCCTCGGGCGTACCGGAGACATTCAGAGTGG 3050  
QY 2402 CCCATCTACAGCGCCCTACTGCTAGCGGCAAGACACTTAAGGTGCGGCTGCTATGCGAG 2461  
Db 3051 CCCATCTACAGCGCCCTACTGCTAGCGGCAAGACACTTAAGGTGCGGCTGCTATGCGAG 3110  
QY 2462 CCCAAGGTTAAGTGTCTTCTGTGAACCGGTGCGGCCACCTAGTGTTCGGGG 2521  
Db 3111 CCCAAGGTTAAGTGTCTTCTGTGAACCGGTGCGGCCACCTAGTGTTCGGGG 3170  
QY 2522 GGTATATGTCTAAGGCACATGCTATCGACCTTAACATCAGAACCGGGTAAAGGACATCA 2581  
Db 3171 GGTATATGTCTAAGGCACATGCTATCGACCTTAACATCAGAACCGGGTAAAGGACATCA 3230  
QY 2582 CCAGGGTGGCCCCCATCACTACTCCACTATGCAAGTTCCTTCCGACCGTGGTTCGT 2641  
Db 3231 CCAGGGTGGCCCCCATCACTACTCCACTATGCAAGTTCCTTCCGACCGTGGTTCGT 3290  
QY 2642 CTGGGGCGCTATGACATCAATAATGATGATGCCACTCAACTGACTCGACCACTA 2701  
Db 3291 CTGGGGCGCTATGACATCAATAATGATGATGCCACTCAACTGACTCGACCACTA 3350  
QY 2702 TCCTGGGACATCGGCACAGTCTTGACACAGCGGAGACGGCTGGAGCGGACTCGTCGTC 2761  
Db 3351 TCCTGGGACATCGGCACAGTCTTGACACAGCGGAGACGGCTGGAGCGGACTCGTCGTC 3410  
QY 2762 TCGCCACCGCTACGCTCCGGGATCGGTCACTGCGCACATCCAAACATCAGAGAGTGG 2821  
Db 3411 TCGCCACCGCTACGCTCCGGGATCGGTCACTGCGCACATCCAAACATCAGAGAGTGG 3470  
QY 2822 CTCTGTCCAGCACTGGAGAAATCCCTTTATGCGAAAGCCATCCCATCGAGACATCA 2881  
Db 3471 CTCTGTCCAGCACTGGAGAAATCCCTTTATGCGAAAGCCATCCCATCGAGACATCA 3530  
QY 2882 AGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAGAAATGTGATGACTCGCGCGA 2941  
Db 3531 AGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAGAAATGTGATGACTCGCGCGA 3590  
QY 2942 AGCTGTCCGGCTCGGACTCAATGTGTAGCATATTACCGGGCTTCATGTATCCGTC 3001  
Db 3591 AGCTGTCCGGCTCGGACTCAATGTGTAGCATATTACCGGGCTTCATGTATCCGTC 3650  
QY 3002 TACCACTAGCGGAGACGTCAATTGCTAGCAACCGGACGTCTTAATGACGGGCTTTACCG 3061  
Db 3651 TACCACTAGCGGAGACGTCAATTGCTAGCAACCGGACGTCTTAATGACGGGCTTTACCG 3710  
QY 3062 GCGATTCGACTGACTGATCGACTGCAATATGTCATGTCACCGACAGTGCACCTCAGCC 3121  
Db 3711 GCGATTCGACTGACTGATCGACTGCAATATGTCATGTCACCGACAGTGCACCTCAGCC 3770  
QY 3122 TGGACCGGACCTTCAACATTTGAGACGACCGGTGCCAAGAGCGGCTGCACGCTCGC 3181  
Db 3771 TGGACCGGACCTTCAACATTTGAGACGACCGGTGCCAAGAGCGGCTGCACGCTCGC 3830  
QY 3182 AGCGGAGGAGGACGACTGTTAGGGGAGGATGGGCAATTAAGTGTTCGATCCAGGAG 3241  
Db 3831 AGCGGAGGAGGAGGACTGTTAGGGGAGGATGGGCAATTAAGTGTTCGATCCAGGAG 3890

QY 3242 AACGGCCCTCGGCGATGTTCCATTCTCGTTCGTTCTGTCGAGTCTATGACGCGGGCTGTG 3301  
Db AAGGGCCCTCGGCGATGTTCCATTCTCGTTCGTTCTGTCGAGTCTATGACGCGGGCTGTG 3950  
QY 3302 CTTGGTACGAGCTCACGCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACAC 3361  
Db CTTGGTACGAGCTCACGCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACAC 4010  
QY 3362 CAGGGTTGCGGCTGTCGCGAGGACCATCTGGAGTTCTGGAGAGGCGCTTTTACAGGCGCTCA 3421  
Db CAGGGTTGCGGCTGTCGCGAGGACCATCTGGAGTTCTGGAGAGGCGCTTTTACAGGCGCTCA 4070  
QY 3422 CCCACATAGACGCCCATTTCTGTCGAGACTAAGCAGGAGGAGAGCAACTTCCCTACC 3481  
Db CCCACATAGACGCCCATTTCTGTCGAGACTAAGCAGGAGGAGAGCAACTTCCCTACC 4130  
QY 3482 TGGTAGCATACCAAGGCTACGGTGTGGGCCAGGGCTCAGGCTCCACCTCOATCGTGGGACC 3541  
Db TGGTAGCATACCAAGGCTACGGTGTGGGCCAGGGCTCAGGCTCCACCTCOATCGTGGGACC 4190  
QY 3542 AAATGTGGAAGTGTCTATACGCGCTAAAGCTTAAGCTAGCTGCACGGCCAAACGCCCTCTGT 3601  
Db AAATGTGGAAGTGTCTATACGCGCTAAAGCTTAAGCTAGCTGCACGGCCAAACGCCCTCTGT 4250  
QY 3602 ATAGGCTGGAGCCGTTCAAAACGAGGTTACTACACACACCCCATAAACCAATATATCA 3661  
Db ATAGGCTGGAGCCGTTCAAAACGAGGTTACTACACACACCCCATAAACCAATATATCA 4310  
QY 3662 TGGCATGCAATGTCCGGCTGACCTCGAGGTGTCTACGAGCACCTGGGTGCTGGTAGGGGAG 3721  
Db TGGCATGCAATGTCCGGCTGACCTCGAGGTGTCTACGAGCACCTGGGTGCTGGTAGGGGAG 4370  
QY 3722 TCTTAGAGCTCTGGCGCGGTATTTGCTGTACAAACAGGAGCGTGTGCTATTTGGGAGGA 3781  
Db TCTTAGAGCTCTGGCGCGGTATTTGCTGTACAAACAGGAGCGTGTGCTATTTGGGAGGA 4430  
QY 3782 TCATCTTGTCCGAAAGCCGCGCATCATTCGCCACAGGGAAGTCTTTACCGGAGTTCG 3841  
Db TCATCTTGTCCGAAAGCCGCGCATCATTCGCCACAGGGAAGTCTTTACCGGAGTTCG 4490  
QY 3842 ATGAGATGGAAGTGTGGCTCTACACCTCCCTTACATCGAAAGGAATTCAGCTCGCGG 3901  
Db ATGAGATGGAAGTGTGGCTCTACACCTCCCTTACATCGAAAGGAATTCAGCTCGCGG 4550  
QY 3902 AACAAATCAAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGCGAGGCTG 3961  
Db AACAAATCAAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGCGAGGCTG 4610  
QY 3962 CTGCTCCGTTGTGGATCCAGTGGGCGGACCTCGAAGGCTTCTGGGCGAAGCATATGT 4021  
Db CTGCTCCGTTGTGGATCCAGTGGGCGGACCTCGAAGGCTTCTGGGCGAAGCATATGT 4670  
QY 4022 GGAATTTTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTCTGCGCTGGCAACCCCG 4081  
Db GGAATTTTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTCTGCGCTGGCAACCCCG 4730  
QY 4082 CGATAGCATCATGTAGGATTCACAGCTCTATCACAGCCCGCTCACACCACCAATATA 4141  
Db CGATAGCATCATGTAGGATTCACAGCTCTATCACAGCCCGCTCACACCACCAATATA 4790  
QY 4142 CCTCTCTGTTTAACTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGGCTG 4201  
Db CCTCTCTGTTTAACTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGGCTG 4850  
QY 4202 CTTCTGCTTTTCGTAGGCGCGCGCATCGCTGGAGCGGCTGTGTGAGCATAGGCTTTGGGA 4261  
Db CTTCTGCTTTTCGTAGGCGCGCGCATCGCTGGAGCGGCTGTGTGAGCATAGGCTTTGGGA 4910  
QY 4262 AGTGCTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCTGCGCT 4321  
Db AGTGCTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCTGCGCT 4970

QY 4322 TTAAAGTCATGATGCGCGAGATGSCCTCCACAGGACCTGGTTAACTACTTCCCTGCTA 4381  
Db TTAAAGTCATGATGCGCGAGATGSCCTCCACAGGACCTGGTTAACTACTTCCCTGCTA 5030  
QY 4382 TCTCTCTCCCTCGCGCCCTAGTCTCGGGGTCTGTGCGCAGCAGATCTGCTGCGGACG 4441  
Db TCTCTCTCCCTCGCGCCCTAGTCTCGGGGTCTGTGCGCAGCAGATCTGCTGCGGACG 5090  
QY 4442 TGGGCCACAGGAGGGGGCTGTGAGTGATGAACCGGCTGATAGGCTTTCGCTTCGCGG 4501  
Db TGGGCCACAGGAGGGGGCTGTGAGTGATGAACCGGCTGATAGGCTTTCGCTTCGCGG 5150  
QY 4502 GTAACCACTCTCCCCACGCACTATGCTGAGAGCAGCGCTGAGCAGCTGCTCACTC 4561  
Db GTAACCACTCTCTCCCCACGCACTATGCTGAGAGCAGCGCTGAGCAGCTGCTCACTC 5210  
QY 4562 AGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAACG 4621  
Db AGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAACG 5270  
QY 4622 AGGACTGCTCAGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGATTTGATATGA 4681  
Db AGGACTGCTCAGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGATTTGATATGA 5330  
QY 4682 CGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGGCTCCTGCCGCGATTCGCCGGAG 4741  
Db CGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGGCTCCTGCCGCGATTCGCCGGAG 5390  
QY 4742 TCCCTCTTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGCGACGGCATCATGC 4801  
Db TCCCTCTTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGCGACGGCATCATGC 5450  
QY 4802 AAACACCTGCCCATGTGAGGACAGATCAACGGACATGTGAATAACGGTTCCATGAGGA 4861  
Db AAACACCTGCCCATGTGAGGACAGATCAACGGACATGTGAATAACGGTTCCATGAGGA 5510  
QY 4862 TGTGGGGCTTAGGACCTGTAGTAAACAGCTGGCATGGAACATTCCTCCATTAACCGGTACA 4921  
Db TGTGGGGCTTAGGACCTGTAGTAAACAGCTGGCATGGAACATTCCTCCATTAACCGGTACA 5570  
QY 4922 CCACGGGCCCTCTGACGCGCTCCCGCGCGCAAAATTTCTAGGGCGCTGTGGCGGCTGG 4981  
Db CCACGGGCCCTCTGACGCGCTCCCGCGCGCAAAATTTCTAGGGCGCTGTGGCGGCTGG 5630  
QY 4982 CTGCTGAGGATGCTGAGGTTACGCGGGTGGGGATTTCCACTAGTGAACGGGATGA 5041  
Db CTGCTGAGGATGCTGAGGTTACGCGGGTGGGGATTTCCACTAGTGAACGGGATGA 5690  
QY 5042 CCACGACAAAGTAAAGTCCCGTGTACAGTTCCGCGCCCGCAATTTCTTACAGAAAGTGG 5101  
Db CCACGACAAAGTAAAGTCCCGTGTACAGTTCCGCGCCCGCAATTTCTTACAGAAAGTGG 5750  
QY 5102 ATGGGTGCGGTTGCAACAGTACGCTCCAGCGTGAACCCCTCTTACGCGGAGGAGTCA 5161  
Db ATGGGTGCGGTTGCAACAGTACGCTCCAGCGTGAACCCCTCTTACGCGGAGGAGTCA 5810  
QY 5162 CATTCCTGTCGGGCTCAATCAATACCTGTTGGGTCAAGCTCCCATGCGAGCCCGAAC 5221  
Db CATTCCTGTCGGGCTCAATCAATACCTGTTGGGTCAAGCTCCCATGCGAGCCCGAAC 5870  
QY 5222 CGGAGCTAGAGTGTCACTTCCATGCTCAGCGCCCTCCACATTTACGGCGGAGACG 5281  
Db CGGAGCTAGAGTGTCACTTCCATGCTCAGCGCCCTCCACATTTACGGCGGAGACG 5930  
QY 5282 CTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGTCCAGCTCATCAGCTAGCCAGC 5341  
Db CTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGTCCAGCTCATCAGCTAGCCAGC 5990  
QY 5342 TGTCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGATCCCGGAGCTGAC 5401  
Db TGTCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGATCCCGGAGCTGAC 6050  
QY 5402 TCATCGAGGCCAACCTCTCTGTGGCGGAGAGTGGCGGGGAACATCACCCGCTGAGT 5461



```
QY 7622 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGACCTGTCTCGTG 7681
Db 8271 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGACCTGTCTCGTG 8330
QY 7682 CCGACCCCGCTGGTTCATGTGGTGGCTACTCCTACTTTCTGTAGGGTAGGCATCTATC 7741
Db 8331 CCGACCCCGCTGGTTCATGTGGTGGCTACTCCTACTTTCTGTAGGGTAGGCATCTATC 8390
QY 7742 TATCCCCCAACCGATGAACCGGGACCTTAAACACTCCAGGCCAATPAGGCCATCTCTGTTTT 7801
Db 8391 TACTCCCCCAACCGATGAACCGGGAGCTTAAACACTCCAGGCCAATPAGGCCATCTCTGTTTT 8450
QY 7802 TTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7861
Db 8451 TTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8510
QY 7862 TTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7921
Db 8511 TTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8570
QY 7922 AGTGTGAAGGTCGGTAGCCGCTTGACTGCAAGAGAGTGTGATAGTGGCTCTCTGCA 7981
Db 8571 AGTGTGAAGGTCGGTAGCCGCTTGACTGCAAGAGAGTGTGATAGTGGCTCTCTGCA 8630
QY 7982 GATCAAGT 7989
Db 8631 GATCAAGT 8638

RESULT 12
US-10-029-907-2
; Sequence 2, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 8642
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
; NAME/KEY: variation
; LOCATION: 6268
; OTHER INFORMATION: r = a or g
; NAME/KEY: variation
; LOCATION: 4446
; OTHER INFORMATION: r = a or g
US-10-029-907-2

Query Match 90.78; Score 7245.2; DB 4; Length 8642;
Best Local Similarity 92.18; Pred. No. 0;
Matches 7965; Conservative 2; Mismatches 10; Indels 675; Gaps 4;

QY 2 CCAGCCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTGT 61
Db 2 CCAGCCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTGT 61
QY 62 CTTACCGCAAGAGCGTCTAGCCATGCGGTAGTATGAGTGTGAGCTCCAGGACC 121
Db 62 CTTACCGCAAGAGCGTCTAGCCATGCGGTAGTATGAGTGTGAGTGTGAGCTCCAGGACC 121
QY 122 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGG 181
```

```
Db 122 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGG 181
QY 182 ACGACCGGGTCCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGCGTCCCGCCG 241
Db 182 ACGACCGGGTCCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGCGTCCCGCCG 241
QY 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTTGGTACTGCTGCTGATAGGG 301
Db 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTTGGTACTGCTGCTGATAGGG 301
QY 302 TGTTCGAGTGCCTCCGGGAGGTCTCGTAGCCGTGCACCATGAGCAGCAATCTTAACC 361
Db 302 TGTTCGAGTGCCTCCGGGAGGTCTCGTAGCCGTGCACCATGAGCAGCAATCTTAACC 361
QY 362 TCAAGAAAAACAAAGGGCGCGCATGATTGAACAAGATGGATTGACACGAGGTTCTCC 421
Db 362 TCAAGAAAAACAAAGGGCGCGCATGATTGAACAAGATGGATTGACACGAGGTTCTCC 421
QY 422 GGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTC 481
Db 422 GGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCTC 481
QY 482 TGATGCGCGCTGTTCGGCTGTGAGCGAGGGCGCGCTTTCTTTTGTCAAGACCGA 541
Db 482 TGATGCGCGCTGTTCGGCTGTGAGCGAGGGCGCGCTTTCTTTTGTCAAGACCGA 541
QY 542 CCTGTCCGTGCTCCTGAATGAATGCAGACGAGGCGCGCTATCGTGGTGGCCAC 601
Db 542 CCTGTCCGTGCTCCTGAATGAATGCAGACGAGGCGCGCTATCGTGGTGGCCAC 601
QY 602 GACGGGCTTCTTGGCAGCTGTGTCGACGTTGTCACCTGAAGCGGGAAGGACTGGCT 661
Db 602 GACGGGCTTCTTGGCAGCTGTGTCGACGTTGTCACCTGAAGCGGGAAGGACTGGCT 661
QY 662 GCTATTGGGCGAAGTSCGGGCGAGATCTCTGTGTCATCTCAGCTTGTCTCCTCCGAGAA 721
Db 662 GCTATTGGGCGAAGTSCGGGCGAGATCTCTGTGTCATCTCAGCTTGTCTCCTCCGAGAA 721
QY 722 AGTATCCATCATGCTGATGCAATGGCGGCTGTCATACGCTTGTGATCGGCTACCTGCC 781
Db 722 AGTATCCATCATGCTGATGCAATGGCGGCTGTCATACGCTTGTGATCGGCTACCTGCC 781
QY 782 ATTTCGACCAACGAGGAAACATCGATCGAGCGAGCACGCTACTCGATGGAAGCGGCT 841
Db 782 ATTTCGACCAACGAGGAAACATCGATCGAGCGAGCACGCTACTCGATGGAAGCGGCT 841
QY 842 TGTTCGATCAGGATGATCTGGAAGAGCATCAGGGCTCGCGCCAGCCGAACTGTTCCG 901
Db 842 TGTTCGATCAGGATGATCTGGAAGAGCATCAGGGCTCGCGCCAGCCGAACTGTTCCG 901
QY 902 CAGGCTCAAGGCGCATGCCGACGAGGATCTGTCGTGACCCCATGCGCATGCTGCTG 961
Db 902 CAGGCTCAAGGCGCATGCCGACGAGGATCTGTCGTGACCCCATGCGCATGCTGCTG 961
QY 962 CTTGCGCAATATCATGTTGGAATGSCCGCTTTCTGTGATTCATGCTGTGSCCGGCT 1021
Db 962 CTTGCGCAATATCATGTTGGAATGSCCGCTTTCTGTGATTCATGCTGTGSCCGGCT 1021
QY 1022 GGGTGTGGGCGCGCTATCAGGACATAGGTTGGCTACCCGTGATATCGCGCTCCCGATTCGCA 1081
Db 1022 GGGTGTGGGCGCGCTATCAGGACATAGGTTGGCTACCCGTGATATCGCGCTCCCGATTCGCA 1081
QY 1082 TGGCGCGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCGCTCCCGATTCGCA 1141
Db 1082 TGGCGCGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCGCTCCCGATTCGCA 1141
QY 1142 GCGCATGCGCTTCTATCGCTTCTTACGAGTTCCTTCTGAGTTCGCGCCAGATGTTAAC 1201
Db 1142 GCGCATGCGCTTCTATCGCTTCTTACGAGTTCCTTCTGAGTTCGCGCCAGATGTTAAC 1201
QY 1190 AGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTA 1249
Db 1190 AGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTA 1249
```

Db 1202 AGACACAACGGTTTCCCTCTAGCGGATCAATTCCG-----CCCCCCCCCTTA 1250  
QY 1250 AGTTTACTGGCCGAAGCCGCTGGATAGAGCCGCTGTCGCTTGTCTATATATATTTT 1309  
Db 1251 AGTTTACTGGCCGAAGCCGCTGGATAGAGCCGCTGTCGCTTGTCTATATATATTTT 1310  
QY 1310 CCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAAACCTGGCCCTGTCTCTTGA 1369  
Db 1311 CCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAAACCTGGCCCTGTCTCTTGA 1370  
QY 1370 CAGCATATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATCAAGTCTGTTGAATGCG 1429  
Db 1371 CAGCATATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATCAAGTCTGTTGAATGCG 1430  
QY 1430 TGAAGGAGCAGTCTCTCTGGAAGCTTCTTGAAGCAAAACACGCTGTGAGGACCCCTT 1489  
Db 1431 TGAAGGAGCAGTCTCTCTGGAAGCTTCTTGAAGCAAAACACGCTGTGAGGACCCCTT 1490  
QY 1490 GCAGGCAGCGAACCCTCCACCTGCGCAGAGTGCTCTGCGGCCAAAGACCACTGTAT 1549  
Db 1491 GCAGGCAGCGAACCCTCCACCTGCGCAGAGTGCTCTGCGGCCAAAGACCACTGTAT 1550  
QY 1550 AAGATACACCTGAAAGCGGCAAAACCCAGTGCCACGTTGTGAGTTGGATGTTGG 1609  
Db 1551 AAGATACACCTGAAAGCGGCAAAACCCAGTGCCACGTTGTGAGTTGGATGTTGG 1610  
QY 1610 AAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGCTGAAGATGCCAGAGG 1669  
Db 1611 AAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGCTGAAGATGCCAGAGG 1670  
QY 1670 TACCCCATTTGATGGATCTGATCTGGGCTCGGTGCACATGCTTTTACATGTGTTAGT 1729  
Db 1671 TACCCCATTTGATGGATCTGATCTGGGCTCGGTGCACATGCTTTTACATGTGTTAGT 1730  
QY 1730 CAGGTTTAAAAACGCTTAGGCCCCCGGAAACCAAGGAGCGTGGTTTCTTTGAAAAAC 1789  
Db 1731 CAGGTTTAAAAACGCTTAGGCCCCCGGAAACCAAGGAGCGTGGTTTCTTTGAAAAAC 1790  
QY 1790 ACGATAATACCATG----- 1803  
Db 1791 ACGATAATACCATGACCGGAGATGGACATCGTGGAGCGCGGTTTTCGTAGTGC 1850  
QY 1804 ----- 1803  
Db 1851 TGATACTCTTGACCTTGTACCGCACTATAAGCTGTTCTCGTAGGCTCATATGTTGT 1910  
QY 1804 ----- 1803  
Db 1911 TACAATATTTATACACAGGCGGAGGACACTTGCAAGTGTGGATCCCGCCCTCAAG 1970  
QY 1804 ----- 1803  
Db 1971 TTCGGGGGGCGGATGCGGTATCTCTCCTCAGTGGCGGATCCACCGAGAGCTAATCT 2030  
QY 1804 ----- 1803  
Db 2031 TTACCAATACCAAAATCTTTGCGCATACTCGGTCCACTCATGTGTCTCAGGCTGTA 2090  
QY 1804 ----- 1803  
Db 2091 TAACCAAGTGGCTATTCTGTCGGCGCACAGGGCTCATTCGTGATGATGCTGGTGC 2150  
QY 1804 ----- 1803  
Db 2151 GGAAGTTGCTGGGGTCAATTATGTCCAAATGGCTCTCATGAAGTTGGCCGCACTGACAG 2210  
QY 1804 ----- 1803  
Db 2211 GTACGTACGTTTATGACCATCTCAOCCCACTGCGGACTGGGCCACGCGGCTACGAG 2270  
QY 1804 ----- 1803  
Db 2271 ACCTTGGGTGGAGTTGAGCCCGTGTCTTCTGTATATGAGACCAAGGTTTACACCT 2330

QY 1804 ----- 1803  
Db 2331 GGGGGGAGACACACCGCGGCTGTGGGACATCATCTTGGCCCTGCCGTCTCGGCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGAGGAGAGATACATCTGGGACCGGACAGACAGCCTTTGAAGGGCAGGGGTGGACCTCC 2450  
QY 1804 --GCGCCTATTACGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCACTA 1861  
Db 2451 TCGCGCCCTATTACGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCACTA 2510  
QY 1862 GCTCACAAGCGCGGACAGGAACAGGTGCGAGGGAGGTCCAAAGTGGTCTCCACGCAA 1921  
Db 2511 GCTCACAAGCGCGGACAGGAACAGGTGCGAGGGAGGTCCAAAGTGGTCTCCACGCAA 2570  
QY 1922 CACAATCTTCTGTGCGACCTGCGTCAATGGCGTGTGTTGAGTCTGTATCATATGTCGG 1981  
Db 2571 CACAATCTTCTGTGCGACCTGCGTCAATGGCGTGTGTTGAGTCTGTATCATATGTCGG 2630  
QY 1982 GCTCAAGACCTTCCCGCGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2041  
Db 2631 GCTCAAGACCTTCCCGCGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2690  
QY 2042 AGGACCTGTCGCTGCGCAAGCGCCCCCGGGCGCGTTCCTTGACACCATGCACCTGCG 2101  
Db 2691 AGGACCTGTCGCTGCGCAAGCGCCCCCGGGCGCGTTCCTTGACACCATGCACCTGCG 2750  
QY 2102 GCAGCTCGGACCTTTACTTGTGTCACGAGCATGCCATGTCTAATCCGTTGCGCGCGCGG 2161  
Db 2751 GCAGCTCGGACCTTTACTTGTGTCACGAGCATGCCATGTCTAATCCGTTGCGCGCGCGG 2810  
QY 2162 GCGACAGCAGGGGAGCCTACTCTCCCCAGGCGCGTCTCCCTACTTGAAGGCTCTTCGG 2221  
Db 2811 GCGACAGCAGGGGAGCCTACTCTCCCCAGGCGCGTCTCCCTACTTGAAGGCTCTTCGG 2870  
QY 2222 GGGTCCACTGTCCTGCCCCCTCGGGCACGCTGTGGGCACTTTTCGGGCTGCGGTGTGA 2281  
Db 2871 GGGTCCACTGTCCTGCCCCCTCGGGCACGCTGTGGGCACTTTTCGGGCTGCGGTGTGA 2930  
QY 2282 CCGGAGGGGTTGCGAAGCGGTGGACTTTGTACCCGTCAGTCTATGGAACCACTATGC 2341  
Db 2931 CCGGAGGGGTTGCGAAGCGGTGGACTTTGTACCCGTCAGTCTATGGAACCACTATGC 2990  
QY 2342 GGTCCCCGGTCTTCAAGGACACTCGTCCCCCTCGGCGCGTACCGGAGACATTTCCAGTGC 2401  
Db 2991 GGTCCCCGGTCTTCAAGGACACTCGTCCCCCTCGGCGCGTACCGGAGACATTTCCAGTGC 3050  
QY 2402 CCATCTTACACGCCCTTACTGTTAGGCGCAAGACACTAAGGTGCGGCTGCGGTATGCAG 2461  
Db 3051 CCATCTTACACGCCCTTACTGTTAGGCGCAAGACACTAAGGTGCGGCTGCGGTATGCAG 3110  
QY 2462 CCGAAGGGTATAAGGTGTTGCTGAAACCCCTGCGTCCGCGCACCCCTAGTGTTCGGGG 2521  
Db 3111 CCGAAGGGTATAAGGTGTTGCTGAAACCCCTGCGTCCGCGCACCCCTAGTGTTCGGGG 3170  
QY 2522 CGTATATGCTTAAGGACATGATGACCTTAAACATCAGAACCGGGGTAAAGACATCA 2581  
Db 3171 CGTATATGCTTAAGGACATGATGACCTTAAACATCAGAACCGGGGTAAAGACATCA 3230  
QY 2582 CCACGGGTGCCCCCATCACTACTCCACTATGCAAGTTCCTTCCGACCGGTGGTTCGT 2641  
Db 3231 CCACGGGTGCCCCCATCACTACTCCACTATGCAAGTTCCTTCCGACCGGTGGTTCGT 3290  
QY 2642 CTGGGGGCGCTTATGACATCAATAATGATGATGCGCACTCACTAAGTCTGACACATA 2701  
Db 3291 CTGGGGGCGCTTATGACATCAATAATGATGATGCGCACTCACTAAGTCTGACACATA 3350  
QY 2702 TCCTGGGATCGGACAGTCTCTGGACCAAGCGGAGACGCTGGAGCGGACCTCGTCTGC 2761  
Db 3351 TCCTGGGATCGGACAGTCTCTGGACCAAGCGGAGACGCTGGAGCGGACCTCGTCTGC 3410



QY 2762 TCGCACCGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTGG 2821  
 Db 3411 TCGCCACCGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTGG 3470  
 QY 2822 CTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCA 2881  
 Db 3471 CTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCA 3530  
 QY 2882 AGGGGGGAGGCACTCATTTTCTGCGCAATTCGAAGAAATGTGATGAGCTCGCGCGGA 2941  
 Db 3531 AGGGGGGAGGCACTCATTTTCTGCGCAATTCGAAGAAATGTGATGAGCTCGCGCGGA 3590  
 QY 2942 AGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTAACCGGGCTTGATGATCGGTCA 3001  
 Db 3591 AGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTAACCGGGCTTGATGATCGGTCA 3650  
 QY 3002 TACCAACTAGCGGAGAGCTCATTTGCTAGCAACGACGCTCTAAAGACGGGCTTTACCG 3061  
 Db 3651 TACCAACTAGCGGAGAGCTCATTTGCTAGCAACGACGCTCTAAAGACGGGCTTTACCG 3710  
 QY 3062 GCGATTTGCACTGATGATCGATGCAATGCAATGATGTCACCCAGACAGTTCAGCC 3121  
 Db 3711 GCGATTTGCACTGATGATGCAATGCAATGATGTCACCCAGACAGTTCAGCC 3770  
 QY 3122 TGGACCGGACTTCAACATGAGACGACGACGTCACCAAGACGCGGTGTACGCTCGC 3181  
 Db 3771 TGGACCGGACTTCAACATGAGACGACGACGTCACCAAGACGCGGTGTACGCTCGC 3830  
 QY 3182 AGCGGGAGGAGGAGCTGGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3241  
 Db 3831 AGCGGGAGGAGGAGCTGGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3890  
 QY 3242 AACGGCCCTCGGGATGTTGCAATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGGCTGTG 3301  
 Db 3891 AACGGCCCTCGGGATGTTGCAATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGGCTGTG 3350  
 QY 3302 CTGTGTTACGAGCTACGCGCGCGGAGACCTCAGTTAGTTGCGGGCTTACTTAAACACAC 3361  
 Db 3951 CTGTGTTACGAGCTACGCGCGCGGAGACCTCAGTTAGTTGCGGGCTTACTTAAACACAC 4010  
 QY 3362 CAGGGTTGCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCGCTCA 3421  
 Db 4011 CAGGGTTGCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCGCTCA 4070  
 QY 3422 CCCACATAGACGCGCAATTTCTGTTCAGACTAAGCAGGACGAGACAACTTCCCTTACC 3481  
 Db 4071 CCCACATAGACGCGCAATTTCTGTTCAGACTAAGCAGGACGAGACAACTTCCCTTACC 4130  
 QY 3482 TGGTAGCATACGAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGACC 3541  
 Db 4131 TGGTAGCATACGAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGACC 4190  
 QY 3542 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGCAAGGCGCCAAACGCGCTGTGT 3601  
 Db 4191 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGCAAGGCGCCAAACGCGCTGTGT 4250  
 QY 3602 ATAGGCTGGAGCGGTTCAAAAAGAGTTACTACCAACACCCCAATACCAATATACATCA 3661  
 Db 4251 ATAGGCTGGAGCGGTTCAAAAAGAGTTACTACCAACACCCCAATACCAATATACATCA 4310  
 QY 3662 TGGCATCATGCTGGCTGACTGGAGTGTGTCAGGACCTGAGTGTGCTGAGGCGGAG 3721  
 Db 4311 TGGCATCATGCTGGCTGACTGGAGTGTGTCAGGACCTGAGTGTGCTGAGGCGGAG 4370  
 QY 3722 TCCTAGCAGCTCTGGCGCGCTATTGCTGCAACAGGACGCTGGTCAATTTGTGGGACGGA 3781  
 Db 4371 TCCTAGCAGCTCTGGCGCGCTATTGCTGCAACAGGACGCTGGTCAATTTGTGGGACGGA 4430  
 QY 3782 TCATCTTGTCCGGAACCGGCCATCATTTCCCGACAGGAGTCTTTACCGGGAGTTCG 3841  
 Db 4431 TCATCTTGTCCGGAACCGGCCATCATTTCCCGACAGGAGTCTTTACCGGGAGTTCG 4490  
 QY 3842 ATGAGATGGAGAGTGGCGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTCGCG 3901

Db 4491 ATGAGATGGAGAGTGGCGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTCGCG 4550  
 QY 3902 AACAAATTTAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTG 3961  
 Db 4551 AACAAATTTAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTG 4610  
 QY 3962 CTGCTCCGCTGTGGAAATCCAGTGGGGACCTCGAAGCCTTCTGGGCAAGCATATGT 4021  
 Db 4611 CTGCTCCGCTGTGGAAATCCAGTGGGGACCTCGAAGCCTTCTGGGCAAGCATATGT 4670  
 QY 4022 GGAATTTTCAATCAGCGGATACAAATTTAGCAGCTTGTCCACTCTGCTGCAACCCCG 4081  
 Db 4671 GGAATTTTCAATCAGCGGATACAAATTTAGCAGCTTGTCCACTCTGCTGCAACCCCG 4730  
 QY 4082 CGATAGCATCACTGATGGCAATTCACAGCTCTATACACAGCCGCTCACACCCCAACATA 4141  
 Db 4731 CGATAGCATCACTGATGGCAATTCACAGCTCTATACACAGCCGCTCACACCCCAACATA 4790  
 QY 4142 CCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTCTTCCAGAGCTG 4201  
 Db 4791 CCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTCTTCCAGAGCTG 4850  
 QY 4202 CTTCTGCTTTCGTAGAGCGCGCATCGCTGAGCGGCTGTGTCAGCATAGGCTTCGGA 4261  
 Db 4851 CTTCTGCTTTCGTAGAGCGCGCATCGCTGAGCGGCTGTGTCAGCATAGGCTTCGGA 4910  
 QY 4262 AGTGTCTGTGGATATTTTGGCAGGTTATGGACAGGGTGGCAGGCGCGCTCGTGGCCT 4321  
 Db 4911 AGTGTCTGTGGATATTTTGGCAGGTTATGGACAGGGTGGCAGGCGCGCTCGTGGCCT 4970  
 QY 4322 TTAAGTCTAGAGCGGGAGATGCCCTCCACCGAGACCTGGTTAACTTACTCCTGCTA 4381  
 Db 4971 TTAAGTCTAGAGCGGGAGATGCCCTCCACCGAGACCTGGTTAACTTACTCCTGCTA 5030  
 QY 4382 TCCTCTCCCTCGCGCTAGTCTGCTGGGCTGTGTGCGCAGCATACTCGCTCGGCACG 4441  
 Db 5031 TCCTCTCCCTCGCGCTAGTCTGCTGGGCTGTGTGCGCAGCATACTCGCTCGGCACG 5090  
 QY 4442 TGGGCCCCAGGGGAGGGGCTGTGCTAGTGGATGAACCGGCTGTAGCTTCGCTTCGCGGG 4501  
 Db 5091 TGGGCCCCAGGGGAGGGGCTGTGCTAGTGGATGAACCGGCTGTAGCTTCGCTTCGCGGG 5150  
 QY 4502 GTAACCAAGCTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCTC 4561  
 Db 5151 GTAACCAAGCTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCTC 5210  
 QY 4562 AGATCTCTCTAGTCTTACCACCTCAGCTGCTGTAAGAGGCTTCAACAGTGGATCAACG 4621  
 Db 5211 AGATCTCTCTAGTCTTACCACCTCAGCTGCTGTAAGAGGCTTCAACAGTGGATCAACG 5270  
 QY 4622 AGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAAGAGATGTTTGGGATTTGATATGCA 4681  
 Db 5271 AGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAAGAGATGTTTGGGATTTGATATGCA 5330  
 QY 4682 CGGTGTGACTGATTTCAAGACCTGCTCCAGTCCAAAGCTTCTGCGCGGCTTCCGCGGAG 4741  
 Db 5331 CGGTGTGACTGATTTCAAGACCTGCTCCAGTCCAAAGCTTCTGCGCGGCTTCCGCGGAG 5390  
 QY 4742 TCCCTCTTCTCATGTCAACGCTGGGTACAAGGGAGTCTTGGCGGGCGACGCGCATCATGC 4801  
 Db 5391 TCCCTCTTCTCATGTCAACGCTGGGTACAAGGGAGTCTTGGCGGGCGACGCGCATCATGC 5450  
 QY 4802 AAACCACTCTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGGA 4861  
 Db 5451 AAACCACTCTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGGA 5510  
 QY 4862 TCGTGGGCGCTAGGACCTGTAGTAAACGCTGGCATGGAACATTTCCCAATTAAACGCGTACA 4921  
 Db 5511 TCGTGGGCGCTAGGACCTGTAGTAAACGCTGGCATGGAACATTTCCCAATTAAACGCGTACA 5570  
 QY 4922 CCACGGGCGCTCAGACGCGCTCCCGGCGCAATTTCTAGGCGCTGCTGCGGGTGG 4981



```

7142 ACCCCACCCGCTTGGCGGCTGCTGGGAGACAGTACACACTCCAGTCAATT 7201
7791 ACCCCACCCGCTTGGCGGCTGCTGGGAGACAGTACACACTCCAGTCAATT 7850
7202 CCTGGCTAGGCAACATCATGATGATCGGCCACCTTGTGGCAAGGATGATCCTGATGA 7261
7851 CCTGGCTAGGCAACATCATGATGATCGGCCACCTTGTGGCAAGGATGATCCTGATGA 7910
7262 CTCATTTCTTCTCCATCTTCTAGCTAGCTCAGGAAACAACTTGAAGGCTAGATGTCAGA 7321
7911 CTCATTTCTTCTCCATCTTCTAGCTAGCTCAGGAAACAACTTGAAGGCTAGATGTCAGA 7970
7322 TCTACGGGGCTGTACTCCATTTAGGACCACTTGACCTACCTCAGATCATTCAGCACTCC 7381
7971 TCTACGGGGCTGTACTCCATTTAGGACCACTTGACCTACCTCAGATCATTCAGCACTCC 8030
7382 ATGGCCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTT 7441
8031 ACGGCCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTT 8090
7442 CATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGATCGGGCCAGAAAGTG 7501
8091 CATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGATCGGGCCAGAAAGTG 8150
7502 TCCGCGCTAGGCTACTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACTCTTCA 7561
8151 TCCGCGCTAGGCTACTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACTCTTCA 8210
7562 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCCCTCCAGTTGGATT 7621
8211 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCCCTCCAGTTGGATT 8270
7622 TATCCAGCTGGTTCGTTGCTGTTTACAGCGGGGAGACATATATCAAGCCTGTCGTG 7681
8271 TATCCAGCTGGTTCGTTGCTGTTTACAGCGGGGAGACATATATCAAGCCTGTCGTG 8330
7682 CCGACCCGCTGCTTCACTGCTGCTGCTACTCTTCTTCTGTTAGGGGTAGGCATCTATC 7741
8331 CCGACCCGCTGCTTCACTGCTGCTGCTACTCTTCTTCTGTTAGGGGTAGGCATCTATC 8390
7742 TACTCCCAACCGATGAACGGGACCTTAAACACTCCAGGCAATAGGCCATCTCTGTTTTT 7801
8391 TACTCCCAACCGATGAACGGGAGCTTAACACTCCAGGCAATAGGCCATCTCTGTTTTT 8450
7802 TTCCC-----TTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCT 7857
8451 TTCCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCC 8510
7858 TTTTTTTTCTCTTTTTTTTCTTTTTTCTTTTTTCTTTTGTGGTCCCATCTTAGCCCTAGTCAC 7917
8511 TTTTTTTTCTCTTTTTTTTCTTTTTTCTTTTTTCTTTTTTGTGGTCCCATCTTAGCCCTAGTCAC 8570
7918 GGCTAGCTGTCAAGGTCCTGAGCCGCTTGACTGCGAGAGAGTCTGATATCTGCGCTCTC 7977
8571 GGCTAGCTGTCAAGGTCCTGAGCCGCTTGACTGCGAGAGAGTCTGATGACTGAGAGTCTGATGACTGCGCTCTC 8630
7978 TGCAGATCAAGT 7989
8631 TGCAGATCAAGT 8642

```

```

RESULT 13
US-10-029-907-6
; Sequence 6, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857

```

```

; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)... (8407)
US-10-029-907-6

Query Match      90.7%; Score 7244; DB 4; Length 8638;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 7957; Conservative 0; Mismatches 20; Indels 671; Gaps 3;

Qy  2  CCAGCCCCCATTTGGGGGGCGACACTCCACCATAGATCACTCCCTCTGTGAGGAACTACTGT 61
Db  2  CCAGCCCCCATTTGGGGGGCGACACTCCACCATAGATCACTCCCTCTGTGAGGAACTACTGT 61

Qy  62  CTTTACGCAAGAAAGCGTCTAGCCATAGCGCTTAGTATGAGTGTGCTGAGCCTCCAGGACC 121
Db  62  CTTTACGCAAGAAAGCGTCTAGCCATAGCGCTTAGTATGAGTGTGCTGAGCCTCCAGGACC 121

Qy  122  CCCCCTCCCGGAGAGACATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGG 181
Db  122  CCCCCTCCCGGAGAGACATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGG 181

Qy  182  ACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTTGGGCGTCCCGCCG 241
Db  182  ACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTTGGGCGTCCCGCCG 241

Qy  242  CGAGACTGTAGCCGAGTGTGTTGGTTCGGAAGAGCCCTTGTGTAAGTACTGCTGATAGGG 301
Db  242  CGAGACTGTAGCCGAGTGTGTTGGTTCGGAAGAGCCCTTGTGTAAGTACTGCTGATAGGG 301

Qy  302  TGTCTGGAGTGTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAC 361
Db  302  TGTCTGGAGTGTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAC 361

Qy  362  TCAAGAAAAACAAAGGGCGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTCC 421
Db  362  TCAAGAAAAACAAAGGGCGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTCC 421

Qy  422  GGCGCTGTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAATTCGCTGCT 481
Db  422  GGCGCTGTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAATTCGCTGCT 481

Qy  482  TGATGCGCGCTGTTCCGGCTGTACGCGAGGGGGCGCCGGTCTTTTCTCAAGACCGA 541
Db  482  TGATGCGCGCTGTTCCGGCTGTACGCGAGGGGGCGCCGGTCTTTTCTCAAGACCGA 541

Qy  542  CCTGTCCGGTGCCTTGAATGAATCTGAGGACGAGGCGCGGCTATTCGCTGGTGGCCAC 601
Db  542  CCTGTCCGGTGCCTTGAATGAATCTGAGGACGAGGCGCGGCTATTCGCTGGTGGCCAC 601

Qy  602  GACGGGCTTCTTGGCAGCTGTGCTGAGCTGCTGAGAGGGGAGGAGGAGTGGCT 661
Db  602  GACGGGCTTCTTGGCAGCTGTGCTGAGCTGCTGAGAGGGGAGGAGGAGTGGCT 661

Qy  662  GCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTGATCTCACTTGTCTCTCCCGAGAA 721
Db  662  GCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTGATCTCACTTGTCTCTCCCGAGAA 721

Qy  722  AGTATCCATCATGCTGATCAATTCGGCGGCTGCAATACGCTTGTATCCGCTACCTGCC 781
Db  722  AGTATCCATCATGCTGATCAATTCGGCGGCTGCAATACGCTTGTATCCGCTACCTGCC 781

Qy  782  ATTGACCAACCAAGCGAAACATCGATTCGAGCGAGCAGTACTCGGATGAAAGCGCGTCT 841
Db  782  ATTGACCAACCAAGCGAAACATCGATTCGAGCGAGCAGTACTCGGATGAAAGCGCGTCT 841

Qy  842  TGTGATCAGGATGATCTGACGCAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCG 901

```

Db	842	TGTGATCAGGATGATCTGGACGAAGAGCATCAGGGCTCGGCAGCCGAATCTTCGC	901
Qy	902	CAGGCTCAAGCGGCGCATGCCGACGGAGAGTCTCGTGACCCATGGCGATGCGTG	961
Db	902	CAGGCTCAAGCGGCGCATGCCGACGGAGAGTCTCGTGACCCATGGCGATGCGTG	961
Qy	962	CTTGCCGAATATCATGTGTGAAATGGCGGCTTTCTGGATTCACTGACTGGCGGCT	1021
Db	962	CTTGCCGAATATCATGTGTGAAATGGCGGCTTTCTGGATTCACTGACTGGCGGCT	1021
Qy	1022	GSGTGTGGCGAACGCTATCAAGACAATAGCGTTGGCTACCGTGATATTCGTAAGAGCT	1081
Db	1022	GSGTGTGGCGAACGCTATCAAGACAATAGCGTTGGCTACCGTGATATTCGTAAGAGCT	1081
Qy	1082	TGGGGCGAATGGGCTGACCGCTTCCCTGCTGTTTAGGATCGCGCTCCCGATTGCA	1141
Db	1082	TGGGGCGAATGGGCTGACCGCTTCCCTGCTGTTTAGGATCGCGCTCCCGATTGCA	1141
Qy	1142	GCGATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTT-----TAAAC	1189
Db	1142	GCGATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTCGCGCCAGATGTTAAC	1201
Qy	1190	AGACCAACAGGTTTCCCTCTAGCGGAGTCAATTCGCGCCCTCTCCCTCCCCCCCCTA	1249
Db	1202	AGACCACAACGTTTCCCTCTAGGGGATCAATTCG-----CCCCCCCCCTA	1250
Qy	1250	ACGTTACTGGCGGAAGCGCTTGGAAATAAGCGCGGTGCGTTGCTATATGTTATTTT	1309
Db	1251	ACGTTACTGGCGGAAGCGCTTGGAAATAAGCGCGGTGCGTTGCTATATGTTATTTT	1310
Qy	1310	CCACATATTCGCTTCTTTGGCAATGTAGGGCGCCGAAACCTGGCCCTGTCTTCTTGA	1369
Db	1311	CCACATATTCGCTTCTTTGGCAATGTAGGGCGCCGAAACCTGGCCCTGTCTTCTTGA	1370
Qy	1370	CGACATTCCTAGGGGTCTTTCCCTCTCGCCAAAGAATGCAAGTCTGTGTAATGTG	1429
Db	1371	CGACATTCCTAGGGGTCTTTCCCTCTCGCCAAAGAATGCAAGTCTGTGTAATGTG	1430
Qy	1430	TGAAGGAGCAGTTCCTCTGAAGCTTCTTGAAGACAACAACGCTGTAGCGACCTTTT	1489
Db	1431	TGAAGGAGCAGTTCCTCTGAAGCTTCTTGAAGACAACAACGCTGTAGCGACCTTTT	1490
Qy	1490	GCAGCAGCGGAACCCGCCACAGTGCCTCTCGGCCAAAAGCCACGCTGAT	1549
Db	1491	GCAGCAGCGGAACCCGCCACAGTGCCTCTCGGCCAAAAGCCACGCTGAT	1550
Qy	1550	AAGTACACCTGMAAGGGGCAACAACCCAGTSCCAGTGTGAGTTGATAGTTGCG	1609
Db	1551	AAGTACACCTGMAAGGGGCAACAACCCAGTSCCAGTGTGAGTTGATAGTTGCG	1610
Qy	1610	AAGAGTCAATAGCTCTCTCAAGCGTATCAACAAGGGGCTGAAGGATGCCAGAAG	1669
Db	1611	AAGAGTCAATAGCTCTCTCAAGCGTATCAACAAGGGGCTGAAGGATGCCAGAAG	1670
Qy	1670	TACCCCATTTGATGGGATCTGATCTGGGGCTCGGTGCAATGTTTTACATGTTTTAGT	1729
Db	1671	TACCCCATTTGATGGGATCTGATCTGGGGCTCGGTGCAATGTTTTACATGTTTTAGT	1730
Qy	1730	CGAGTTTAAAAACGCTAGCCCCCGAACCAAGGGGAGCGTGGTTTTCTTTGAAAAAC	1789
Db	1731	CGAGTTTAAAAACGCTAGCCCCCGAACCAAGGGGAGCGTGGTTTTCTTTGAAAAAC	1790
Qy	1790	ACGATAATPACCATG-----	1803
Db	1791	ACGATAATPACCATGACCGGAGATGCGAGATCGTCGGAGGCGCGGTTTTCGTAGGTC	1850
Qy	1804	-----	1803
Db	1851	TGATACTCTTGACTACCGCATATTAAGCTGTTTCTCGCTAGSCTCATATGGTGT	1910
Qy	1804	-----	1803
Db	1911	TACAATATTTTATCACAGGGCGAGGACACTTGCAAGTGTGATCCCCCCCTCAACG	1971
Qy	1804	-----	1803
Db	1971	TTCGGGGGGGCGCGATGCGGTATCTCTCTACGTGCGGATCCACCAGAGCTTAATCT	2030
Qy	1804	-----	1803
Db	2031	TTACCATCACAAAATCTTGCTCGCCATACTCGGTCCACTCATGTGTGCTCCAGGCTGTA	2090
Qy	1804	-----	1803
Db	2091	TAAACAAAGTGCCTACTTCGTGCGGCGACACGGGCTCATTCGTGATGATGCTGTGTC	2150
Qy	1804	-----	1803
Db	2151	GGAAGTTGCTGGGGGTCAATTATGTCCAATGGCTCTCATGAAGTTGGCGGCACTGACAG	2210
Qy	1804	-----	1803
Db	2211	GTAGTACGTTTATGACCATCTCACCCACTCGCGGACTGGGCCACGCGGCGCTACGAG	2270
Qy	1804	-----	1803
Db	2271	ACCTTGGGTGGCAGTTGAGCCGCTCTTCTCTGATATGGAGACCAAAGTTTATCACCT	2330
Qy	1804	-----	1803
Db	2331	GGGGGCGACACCGCGCGGTGTGGGACATCATCTTGGGCGTCCCGCTCTCGCCCGCA	2390
Qy	1804	-----	1803
Db	2391	GGGGGAGGAGATACATCTGGGACCGGACAGACGCTTGAAGGCGAGGGTGGGACTCC	2450
Qy	1804	--CGGCTATTACGGCTACTCCCAAACAGACGGAGGCTACTTGGCTGCATCATCACTA	1861
Db	2451	TCGCGCTATTACGGCTACTCCCAAACAGACGGAGGCTACTTGGCTGCATCATCACTA	2510
Qy	1862	GCCTCAGCGCGGACAGAACACAGTCTGAGGGGAGTCCAAAGTGTCTCCACCGAA	1921
Db	2511	GCCTCAGCGCGGACAGAACACAGTCTGAGGGGAGTCCAAAGTGTCTCCACCGAA	2570
Qy	1922	CACAACTTTCTTGGCGACTGCTCAATGGGCTGTGTGACTGTCTATCATGTGTCCG	1981
Db	2571	CACAACTTTCTTGGCGACTGCTCAATGGGCTGTGTGACTGTCTATCATGTGTCCG	2630
Qy	1982	GCTCAAAGACCTTTGCCGCCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2041
Db	2631	GCTCAAAGACCTTTGCCGCCCAAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2690
Qy	2042	AGGACCTCTGCTGGGAGCGCCCCCGGGCGGTTCTTGACACCATGCACTGCG	2101
Db	2691	AGGACCTCTGCTGGGAGCGCCCCCGGGCGGTTCTTGACACCATGCACTGCG	2750
Qy	2102	GCACTCTGGAACCTTTACTTGGTCAAGGACATGCCGATGCTCATTCGGTGCSCCGCGGG	2161
Db	2751	GCACTCTGGAACCTTTACTTGGTCAAGGACATGCCGATGCTCATTCGGTGCSCCGCGGG	2810
Qy	2162	GCGACAGCAGGGGAGCTTCTCTCCCCAGCGCGGTCTCTTACCTTGAAGGGCTCTTGG	2221
Db	2811	GCGACAGGGGAGGAGCTTCTCTCCCCAGCGCGGTCTCTTACCTTGAAGGGCTCTTGG	2870
Qy	2222	GSGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTCGCGGTGCA	2281
Db	2871	GSGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTCGCGGTGCA	2930
Qy	2282	CCGAGGGGTTCGAAGCGGTGGACTTTGTACCGCTCGAGTCTATGGAACCACTATGC	2341
Db	2931	CCGAGGGGTTCGAAGCGGTGGACTTTGTACCGCTCGAG	



Db	5211	AGATCCCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACATGATCAACG	5270
Qy	4622	AGGACTGCTCCACGCCATGCTCCGGCTCGCTGGCTAAGAGATGTTTGGGATTGGATATGCA	4681
Db	5271	AGGACTGCTCCACGCCATGCTCCGGCTCGCTGGCTAAGAGATGTTTGGGATTGGATATGCA	5330
Qy	4682	CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCCTGCGCGGATTCGCCGGAG	4741
Db	5331	CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCCTGCGCGGATTCGCCGGAG	5390
Qy	4742	TCCCTTCTTCTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGGCGACGGCATCATGC	4801
Db	5391	TCCCTTCTTCTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGGCGACGGCATCATGC	5450
Qy	4802	AAACCACTGCCCATGTGGACACAGATACCCGACATGTGAATAAGGTTTCCATCAGGA	4861
Db	5451	AAACCACTGCCCATGTGGACACAGATACCCGACATGTGAATAAGGTTTCCATCAGGA	5510
Qy	4862	TCGTGGGGCTTAGGACCTGTAGTAAACGTGGCATGGAACTTCCCATTTAAACGCGTACA	4921
Db	5511	TCGTGGGGCTTAGGACCTGTAGTAAACGTGGCATGGAACTTCCCATTTAAACGCGTACA	5570
Qy	4922	CCACGGGCCCCGTGACCGCCCTCCCGGGCGCAATTAATTCTAGGGCGCTGTGCGGGTGG	4981
Db	5571	CCACGGGCCCCGTGACCGCCCTCCCGGGCGCAATTAATTCTAGGGCGCTGTGCGGGTGG	5630
Qy	4982	CTGCTGAGGAGTAGTGGAGGTTACCGGGGTGGGGATTTCCACTACGTGACGGGCATGA	5041
Db	5631	CTGCTGAGGAGTAGTGGAGGTTACCGGGGTGGGGATTTCCACTACGTGACGGGCATGA	5690
Qy	5042	CCACTGACAAAGTAAAGTGCCTGTTCAGGTTCGGGCCCCCGAAATTTCTTACAGAAAGTGG	5101
Db	5691	CCACTGACAAAGTAAAGTGCCTGTTCAGGTTCGGGCCCCCGAAATTTCTTACAGAAAGTGG	5750
Qy	5102	ATGGGGTGGGGTTGACACAGTACGCTCCAGCTGCAAAACCCCTCTTACGGGAGGAGTCA	5161
Db	5751	ATGGGGTGGGGTTGACACAGTACGCTCCAGCTGCAAAACCCCTCTTACGGGAGGAGTCA	5810
Qy	5162	CAATTCCTGGTCGGGCTCAATCAATACCTGGTTCGGGTACAGCTCCCATCGGAGCCCGAAC	5221
Db	5811	CAATTCCTGGTCGGGCTCAATCAATACCTGGTTCGGGTACAGCTCCCATCGGAGCCCGAAC	5870
Qy	5222	CGGACGTAGCAGTGTCTCACTTCCATGCTCACCCAGCCCTCCCACTATACGGCGGAGACGG	5281
Db	5871	CGGACGTAGCAGTGTCTCACTTCCATGCTCACCCAGCCCTCCCACTATACGGCGGAGACGG	5930
Qy	5282	CTAAGCGTAGGCTGGCGAGGGATCTCCGCCCTCTCTTGGCGAGCTCATCAGCTAGCCAGC	5341
Db	5931	CTAAGCGTAGGCTGGCGAGGGATCTCCGCCCTCTCTTGGCGAGCTCATCAGCTAGCCAGC	5990
Qy	5342	TGTCCTGGCTTCTTGAAGGCAATGACTACCCGTCTATGACTCCCGGAGCGCTGACC	5401
Db	5991	TGTCCTGGCTTCTTGAAGGCAATGACTACCCGTCTATGACTCCCGGAGCGCTGACC	6050
Qy	5402	TCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGGAACATCACCCGCTGGAGT	5461
Db	6051	TCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGGAACATCACCCGCTGGAGT	6110
Qy	5462	CAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTTCAAGCGGAGGAGTGA	5521
Db	6111	CAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTTCAAGCGGAGGAGTGA	6170
Qy	5522	GGGAGGTATCCGTTCCGGCGGAGATCCTCGGAGTCCAGGAAATTTCCCTCGAGCGATGC	5581
Db	6171	GGGAGGTATCCGTTCCGGCGGAGATCCTCGGAGTCCAGGAAATTTCCCTCGAGCGATGC	6230
Qy	5582	CCATATGGGCACCCCGGATTACAAACCTCCACTGTTAGTCTCTGGAAGGACCCGGACT	5641
Db	6231	CCATATGGGCACCCCGGATTACAAACCTCCACTGTTAGTCTCTGGAAGGACCCGGACT	6290
Qy	5642	ACGTCCCTCCAGTGGTACACGGGTGTCATGTGCGGCTGCGCAAGGCCCTCCGATACCAAC	5701

Db	6291	ACGTCCCTCCASITGGTATACAGGGGTGTCCACTGCCCGCTGCCTCAAGAGCCGCTCCGATACCAC	6335
Qy	5702	CTCCACGAGAGAGAGACGGTTGTCTGTCAAGATCTACCGTGTCTTCTGCTTCGGCGG	5761
Db	6351	CTCCACGAGAGAGAGACGGTTGTCTGTCAAGATCTACCGTGTCTTCTGCTTCGGCGG	6410
Qy	5762	AGCTCGCCACAAAGACCTTCGGCAGCTCCGAATCGTCGGCCGTGCACAGCGGCACGGCAA	5821
Db	6411	AGCTCGCCACAAAGACCTTCGGCAGCTCCGAATCGTCGGCCGTGCACAGCGGCACGGCAA	6470
Qy	5822	CGGCTCTCTCTGACACGCCCTCCGACGACGGCGACGGGATCCGACGTTGAGTCGTACT	5881
Db	6471	CGGCTCTCTCTGACACGCCCTCCGACGACGGCGACGGGATCCGACGTTGAGTCGTACT	6530
Qy	5882	CCTCCATGCCCCCTTGTAGGGGAGCGGGGATCCCAGATCTCAGGACGGGTCTTGGT	5941
Db	6531	CCTCCATGCCCCCTTGTAGGGGAGCGGGGATCCCAGATCTCAGGACGGGTCTTGGT	6590
Qy	5942	CTACCGTAAGCGAGGAGGCTAGTGAGGACGTCGCTGCTGCTCGATGTCTACACATGGA	6001
Db	6591	CTACCGTAAGCGAGGAGGCTAGTGAGGACGTCGCTGCTGCTCGATGTCTACACATGGA	6650
Qy	6002	CAGCGGCCCTGATCAACGCCATGGCTGCGGAGAAACCAAGCTGCCCATCAATGCATGA	6061
Db	6651	CAGCGGCCCTGATCAACGCCATGGCTGCGGAGAAACCAAGCTGCCCATCAATGCATGA	6710
Qy	6062	GCAACTCTTTGCTCCGTCACCAAACTTTGGTCTATGCTACACATCTCCGACGGCGAAGCC	6121
Db	6711	GCAACTCTTTGCTCCGTCACCAAACTTTGGTCTATGCTACACATCTCCGACGGCGAAGCC	6770
Qy	6122	TGCGGCAGAAAGAGGTCAACCTTTGACAGACTGCAGGTCTTGAGACGACCACTACCGGACG	6181
Db	6771	TGCGGCAGAAAGAGGTCAACCTTTGACAGACTGCAGGTCTTGAGACGACCACTACCGGACG	6830
Qy	6182	TGCTCAAGGAGTAGAAGCGAAGCGGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAGG	6241
Db	6831	TGCTCAAGGAGTAGAAGCGAAGCGGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAGG	6890
Qy	6242	AAGCCTGTAGCTGACGCCGCCACATTCGGCCAGATCTAAATTTGGCTATGGGCGAAGG	6301
Db	6891	AAGCCTGTAGCTGACGCCGCCACATTCGGGCCAGATCTAAATTTGGCTATGGGCGAAGG	6950
Qy	6302	ACGTCCGGAACCTATCCAGCAAGCGCGTTAAACACATCCGCTCCGTGTGGAAGGACTTGC	6361
Db	6951	ACGTCCGGAACCTATCCAGCAAGCGCGTTAAACACATCCGCTCCGTGTGGAAGGACTTGC	7010
Qy	6362	TGGAAGACACTGAGACACCAATTTGACACCACTCATGCGCAAAAAATGAGTTTCTGCG	6421
Db	7011	TGGAAGACACTGAGACACCAATTTGACACCACTCATGCGCAAAAAATGAGTTTCTGCG	7070
Qy	6422	TCCAAACAGAGAGGGGGCGGAAGCGAGCTCGCCTTATCGTATTCAGATTTGGGGG	6481
Db	7071	TCCAAACAGAGAGGGGGCGGAAGCGAGCTCGCCTTATCGTATTTCCAGATTTGGGGG	7130
Qy	6482	TTCTGTGTGCGAGAAAATGCGCCTTTACGATGTGTCTCCACCCCTCCCTCAGGCGCTGA	6541
Db	7131	TTCTGTGTGCGAGAAAATGCGCCTTTACGATGTGTCTCCACCCCTCCCTCAGGCGCTGA	7190
Qy	6542	TGGGCTCTTCATACGGATTCCAATCTCTCTGGACAGCGGGTTCGAGTTCTCTGGTGAATG	6601
Db	7191	TGGGCTCTTCATACGGATTCCAATCTCTCTGGACAGCGGGTTCGAGTTCTCTGGTGAATG	7250
Qy	6602	CCTGGAAGCGAAGATGCCCCATATGGGCTTCGCATATGACACCCGCTGTTTGTACTCAA	6661
Db	7251	CCTGGAAGCGAAGATGCCCCATATGGGCTTCGCATATGACACCCGCTGTTTGTACTCAA	7310
Qy	6662	CGGTCACTGAGAAATGACATCCGTGTTGAGAGTCAATCTACCAATGTTGTGACTTGGCCC	6721
Db	7311	CGGTCACTGAGAAATGACATCCGTGTTGAGAGTCAATCTACCAATGTTGTGACTTGGCCC	7370
Qy	6722	CCGAAGCCAGACAGGCAATAGGTTCGCTCACAGAGCGGTTTACATCGGGGGCCCCCTGA	6781
Db	7371	CCGAAGCCAGACAGGCAATAGGTTCGCTCACAGAGCGGTTTACATCGGGGGCCCCCTGA	7430





QY 481 CTGATGCGCGCTGTTCGGCTGTCTCAGCGAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540  
DB |||||  
DB 481 CTGATGCGCGCTGTTCGGCTGTCTCAGCGAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540  
QY 541 ACTGTTCGGGTGCGCTGTGAATGAATCTGACGAGAGCGAGCGCGCTATCGTGGCTGGCCA 600  
DB |||||  
DB 541 ACTGTTCGGGTGCGCTGTGAATGAATCTGACGAGAGCGAGCGCGCTATCGTGGCTGGCCA 600  
QY 601 CGACGGGCGCTTCTCTTCGCGAGCTGTCTGCTGAGCTGTCTCACTGAAGCGGAAGGACTGGC 660  
DB |||||  
DB 601 CGACGGGCGCTTCTCTTCGCGAGCTGTCTGCTGAGCTGTCTCACTGAAGCGGAAGGACTGGC 660  
QY 661 TCTATTTGGCGGAAGTGCAGGGGAGGATCTCTCTGTCACTCTCACCTTGTCTCTGCGGAGA 720  
DB |||||  
DB 661 TCTATTTGGCGGAAGTGCAGGGGAGGATCTCTCTGTCACTCTCACCTTGTCTCTGCGGAGA 720  
QY 721 AAGTATCCATCATGCTGTGATGCAATGCGGGGCTGATACGCTTGTATCGCGCTACCTGCC 780  
DB |||||  
DB 721 AAGTATCCATCATGCTGTGATGCAATGCGGGGCTGATACGCTTGTATCGCGCTACCTGCC 780  
QY 781 CATTCGACCAACGAGCAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC 840  
DB |||||  
DB 781 CATTCGACCAACGAGCAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGTC 840  
QY 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCCAGCGCACTTTCCG 900  
DB |||||  
DB 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCCAGCGCACTTTTCG 900  
QY 901 CAGGCTCAAGGCGCGATGCGCGAGCGAGATCTCGTGTGACCCATGCGGATGCGCT 960  
DB |||||  
DB 901 CAGGCTCAAGGCGCGATGCGCGAGCGAGATCTCGTGTGACCCATGCGGATGCGCT 960  
QY 961 GCTTCCGGAATCATGCTGGAATAAGCGCGCTTTCTGATTCATCGACTGTGGCCGCG 1020  
DB |||||  
DB 961 GCTTCCGGAATCATGCTGGAATAAGCGCGCTTTCTGATTCATCGACTGTGGCCGCG 1020  
QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGC 1080  
DB |||||  
DB 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGC 1080  
QY 1081 TTGGGGGCAATGCGGTGACCGCTTCTCTGCTGTTTACGGTATCGCGCTCCGATTCGC 1140  
DB |||||  
DB 1081 TTGGGGGCAATGCGGTGACCGCTTCTCTGCTGTTTACGGTATCGCGCTCCGATTCGC 1140  
QY 1141 AGCGATCGCTTCTATCGCTTCTTACGAGTTCTTCTGAGTT-----TAAA 1188  
DB |||||  
DB 1141 AGCGATCGCTTCTATCGCTTCTTACGAGTTCTTCTGAGTTCTGAGTTCTGCGCCGAGATTTAA 1200  
QY 1189 CAGACCAACCGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248  
DB |||||  
DB 1201 CAGACCAACCGTTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCT 1249  
QY 1249 AAGCTTACTGGCGAAGCGCTTGGAAATAGGCGGCTGTGGTTTGTCTATATGTTATTT 1308  
DB |||||  
DB 1250 AAGCTTACTGGCGAAGCGCTTGGAAATAGGCGGCTGTGGTTTGTCTATATGTTATTT 1309  
QY 1309 TCCACCATATGCGCTTTTGGCAATGTAGGCGCCGAAACCTGCGCTGCTCTTTG 1368  
DB |||||  
DB 1310 TCCACCATATGCGCTTTTGGCAATGTAGGCGCCGAAACCTGCGCTGCTCTTTG 1369  
QY 1369 ACGAGCATTCCTAGGGGTCTTCCCTCTCGCAAGGAAATGCAAGGTCTGTTGAATGTC 1428  
DB |||||  
DB 1370 ACGAGCATTCCTAGGGGTCTTCCCTCTCGCAAGGAAATGCAAGGTCTGTTGAATGTC 1429  
QY 1429 GTGAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTT 1488  
DB |||||  
DB 1430 GTGAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTT 1489  
QY 1489 TCGAGCAGCGGAACCCCGACCTGCGGACAGTGTCTTGGGCGCAAAAGCCAGCTGTA 1548  
DB |||||  
DB 1490 TCGAGCAGCGGAACCCCGACCTGCGGACAGTGTCTTGGGCGCAAAAGCCAGCTGTA 1549  
QY 1549 TAAGATACCTTGCAAAAGCGGCAACACCCAGTGCCACGTTGTGATGGATGTTGTG 1608

DB 1550 TAAGATACCTTGCAAAAGCGGCAACACCCAGTGCCAGTGTGTGATGGATGTTGT 1609  
QY |||||  
QY 1609 GAAAGAGTCAATGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAAG 1668  
DB |||||  
DB 1610 GAAAGAGTCAATGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAAG 1669  
QY |||||  
QY 1669 GTACCCCATTTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTGTTAG 1728  
DB |||||  
DB 1670 GTACCCCATTTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTGTTAG 1729  
QY |||||  
QY 1729 TCGAGTTAAAAAAGCTGTAGGCCCGCGAAACAAGGGAAGCTGTTTCTTTGAAAAA 1788  
DB |||||  
DB 1730 TCGAGTTAAAAAAGCTGTAGGCCCGCGAAACAAGGGAAGCTGTTTCTTTGAAAAA 1789  
QY |||||  
QY 1789 CAGGATATACCATG----- 1803  
DB |||||  
DB 1790 CAGGATATACCATGGAACGGGAGATGGCAGCATCGTGGAGGCGCGTTCGTAGGT 1849  
QY |||||  
QY 1804 ----- 1803  
DB 1850 CTGATACTCTTGACCTTGTACCGGCACTATAAGCTGTTCTCGTAGGCTCATATGTTG 1909  
QY |||||  
QY 1804 ----- 1803  
DB 1910 TTACAATATTTTATCACAGGCGGAGGCACACTTGCAAGTGTGGATCCCCCCTCAAC 1969  
QY |||||  
QY 1804 ----- 1803  
DB 1970 GTTCGGGGGCGCGGATGCGCTCATCTCTCAAGTGTGGCGGATCCACAGAGCTAATC 2029  
QY |||||  
QY 1804 ----- 1803  
DB 2030 TTACCAATACCAAAATCTTGCTCGCCATCTCGGTCCAATCATGTGTCTCCAGGCTGT 2089  
QY |||||  
QY 1804 ----- 1803  
DB 2090 ATACCAAAAGTGGCTACTCTGTCGCGCACACGGGCTCATCTGTCATGATGTCATG 2149  
QY |||||  
QY 1804 ----- 1803  
DB 2150 CGGAAGTTGTGGGGTCAATTATGTCCAAATGGCTCTCATGAAGTTGGCGGCACTGACA 2209  
QY |||||  
QY 1804 ----- 1803  
DB 2210 GGTACGTAAGTTATGACCATCTCACCCACTCTCGGACTCGGGCCACGCGGCGCTACGA 2269  
QY |||||  
QY 1804 ----- 1803  
DB 2270 GACCTTGGCGTGCGAGTTGAGCCCGTCTCTCTGATATGAGAGCAAGGTTATCAC 2329  
QY |||||  
QY 1804 ----- 1803  
DB 2330 TGGGGGAGAGACACCGCGCGCTGTGGGGGACATCATCTTGGGCTGCGCTCTCCGCCCGC 2389  
QY |||||  
QY 1804 ----- 1803  
DB 2390 AGGGGAGGAGATACATCTGGGACCGGACAGACCTTGAAGGGCAGGGTGGCGACTC 2449  
QY |||||  
QY 1804 ---GGGCTTATTACGGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCAATCATCAT 1860  
DB |||||  
DB 2450 CTCGGGCTTATTACGGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCAATCATCAT 2509  
QY |||||  
QY 1861 AGCCTCACAGGCGGACAGGAAACAGGTCGAGGGGAGTCCAGTGGTCTCCACCGCA 1920  
DB |||||  
DB 2510 AGCCTCACAGGCGGACAGGAAACAGGTCGAGGGGAGTCCAGTGGTCTCCACCGCA 2569  
QY |||||  
QY 1921 ACACATCTTCTCGCGACCTGCGTCAATGCGGTGTGTGGACTGTCTTATCATGTGTC 1980  
DB |||||  
DB 2570 ACACATCTTCTCGCGACCTGCGTCAATGCGGTGTGTGGACTGTCTTATCATGTGTC 2629  
QY |||||  
QY 1981 GGCTCAAGACCTTCTCGCGCGGCGAAGGGCGCAATCACCAANTGTACCAANTGTGGAC 2040  
DB |||||

Db 2630 GGCTCAAAGACCTTGGCGGCCCAAGGGCCCAATACACCCAAATGTACACCAATGTGGAC 2689  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCGCCCGGGGCGGTTCTTGTGAACCAATGCACTGCG 2100  
Db 2690 CAGGACCTCGTGGCTGGCAAGCGCCGCCCGGGGCGGTTCTTGTGAACCAATGCACTGCG 2749  
QY 2101 GGCAGCTCGGACCTTTACTTGGTTCACGAGCATGCGCATGTCATTCGCGTGGCGCGCGG 2160  
Db 2750 GGCAGCTCGGACCTTTACTTGGTTCACGAGCATGCGCATGTCATTCGCGTGGCGCGCGG 2809  
QY 2161 GGCAGACAGCGGGAGCCTACTCTCCGCCAGGCGCGTCTCTACTTTGAAGGGCTCTTCG 2220  
Db 2810 GGCAGACAGCGGGAGCCTACTCTCCGCCAGGCGCGTCTCTACTTTGAAGGGCTCTTCG 2869  
QY 2221 GCGGTTCACTGCTCTGCGCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGCG 2280  
Db 2870 GCGGTTCACTGCTCTGCGCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGCG 2929  
QY 2281 ACCCGAGGGTTGGAAAGGGGTGGACTTTGTACCCGTCGAGTCTATGGAACACCACTATG 2340  
Db 2930 ACCCGGGGGTTGGAAAGGGGTGGACTTTGTACCCGTCGAGTCTATGGAACACCACTATG 2989  
QY 2341 CGTCCCGGCTTTACGGCAAACTCGTCCCTCCGCGCGTACCGCAGACATTCAGGGT 2400  
Db 2990 CGTCCCGGCTTTACGGCAAACTCGTCCCTCCGCGCGTACCGCAGACATTCAGGGT 3049  
QY 2401 GCCCATCTACGCCCTACTGTTAGGGGCAAGAGCACTAAGTTCGGGCTGCGTATGCA 2460  
Db 3050 GCCCATCTACGCCCTACTGTTAGGGGCAAGAGCACTAAGTTCGGGCTGCGTATGCA 3109  
QY 2461 GCCCAAGGTTAAGGTGCTTCTCAACCCGTCGTCGCGCCACCCCTAGGTTTCGGG 2520  
Db 3110 GCCCAAGGTTAAGGTGCTTCTCAACCCGTCGTCGCGCCACCCCTAGGTTTCGGG 3169  
QY 2521 GCGTATATGTAAAGCACATGTTATCGACCCCTAAACATCAAGACCGGGGTAAGGACCATC 2580  
Db 3170 GCGTATATGTAAAGCACATGTTATCGACCCCTAAACATCAAGACCGGGGTAAGGACCATC 3229  
QY 2581 ACCACGGTGCGCCCTACGTTACTCAACCTATGGCAAGTTTCTTGGCGAGCGGTGTC 2640  
Db 3230 ACCACGGTGCGCCCTACGTTACTCAACCTATGGCAAGTTTCTTGGCGAGCGGTGTC 3289  
QY 2641 TCTGGGGCGCTATGACATCAATAATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db 3290 TCTGGGGCGCTATGACATCAATAATGTGATGAGTGCCACTCAACTGACTCGACCACT 3349  
QY 2701 ATCTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGCGTGAGCGCGCACTCGTCGTG 2760  
Db 3350 ATCTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGCGTGAGCGCGCACTCGTCGTG 3409  
QY 2761 CTGCCACCGCTAGCGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTG 2820  
Db 3410 CTGCCACCGCTAGCGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTG 3469  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCCATCGAGACCATC 2880  
Db 3470 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCCATCGAGACCATC 3529  
QY 2881 AAGGGGGGAGGACCTCATTTTCTGCAATTCGAAGAAATGTGATGAGTCTGCGCGG 2940  
Db 3530 AAGGGGGGAGGACCTCATTTTCTGCAATTCGAAGAAATGTGATGAGTCTGCGCGG 3589  
QY 2941 AAGTGTCCGCGCTCGGACCTCAATGTGTAGCATATTACGGGGCTTGTGATTCGCTC 3000  
Db 3590 AAGTGTTCGCGCTCGGACCTCAATGTGTAGCATATTACGGGGCTTGTGATTCGCTC 3049  
QY 3001 ATACCAACTAGCGAGACGTCATTGTCTGAGCAACGCGCTCTTAATGAGCGGGTTTACC 3060  
Db 3650 ATACCAACTAGCGAGACGTCATTGTCTGAGCAACGCGCTCTTAATGAGCGGGTTTACC 3709  
QY 3061 GGGCAATTCGACTCAGTGATCGACTGCAATATACATGTGTCAACGAGACAGTTCAGC 3120  
Db 3710 GGGCAATTCGACTCAGTGATCGACTGCAATATACATGTGTCAACGAGACAGTTCAGC 3769

QY 3121 CTGGAACCGGACCTTCACTTTGAGACGACACCGTGCACAGACGCGGTGTGACGCTCG 3180  
Db 3770 CTGGAACCGGACCTTCACTTTGAGACGACACCGTGCACAGACGCGGTGTGACGCTCG 3829  
QY 3181 CAGCGGCGAGGACGACTGCTAGGGGAGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3830 CAGCGGCGAGGACGACTGCTAGGGGAGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3889  
QY 3241 GAAACGCGCTCGGGCAATGTTCTCGATTCTCTGCGAGTGTCTATGACGCGGGTGT 3300  
Db 3890 GAAACGCGCTCGGGCAATGTTCTCGATTCTCTGCGAGTGTCTATGACGCGGGTGT 3949  
QY 3301 GCTTGTGTACAGCTCAGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTTAAACACA 3360  
Db 3950 GCTTGTGTACAGCTCAGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTTAAACACA 4009  
QY 3361 CAGGGTTGCGCGTCTGCCAGGACCATCTCGAGTCTTGGGAGAGCGTCTTTACAGGGCTC 3420  
Db 4010 CAGGGTTGCGCGTCTGCCAGGACCATCTCGAGTCTTGGGAGAGCGTCTTTACAGGGCTC 4069  
QY 3421 ACCACATAGACGCCAATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 3480  
Db 4070 ACCACATAGACGCCAATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 4129  
QY 3481 CTGGTAGCATACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
Db 4130 CTGGTAGCATACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCATCGTGGGAC 4189  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTGACGCTGACAGGGGCAACGCGCTCGT 3600  
Db 4190 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACAGGGGCAACGCGCTCGT 4249  
QY 3601 TATAGGCTGGAGCGGTTTCAAAAAGAGTTTACTACCAACACCCCAATPAAACCAATACATC 3660  
Db 4250 TATAGGCTGGAGCGGTTTCAAAAAGAGTTTACTACCAACACCCCAATPAAACCAATACATC 4309  
QY 3661 ATGGCATGATGTGCGTGTGAGTGTGCTCAGAGACCTGGGTGCTGGTAGGGGA 3720  
Db 4310 ATGGCATGATGTGCGTGTGAGTGTGCTCAGAGACCTGGGTGCTGGTAGGGGA 4369  
QY 3721 GTCTAGACGCTGTGGCGGTTATGCTGCAACAGGAGCGTGTGTCATTTGCGGGCAGG 3780  
Db 4370 GTCTAGACGCTGTGGCGGTTATGCTGCAACAGGAGCGTGTGTCATTTGCGGGCAGG 4429  
QY 3781 ATCATCTTGTCCGGAAGCGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 4430 ATCATCTTGTCCGGAAGCGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 4489  
QY 3841 GATGAGATGGAAGTGTGCGCTCAGCTTCTTATCATCGAACAGGGAATGCACTCGCC 3900  
Db 4490 GATGAGATGGAAGTGTGCGCTCAGCTTCTTATCATCGAACAGGGAATGCACTCGCC 4549  
QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGT 3960  
Db 4550 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGT 4609  
QY 3961 GCTGCTCCCGTGTGGAATCCAAAGTGTGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
Db 4610 GCTGCTCCCGTGTGGAATCCAAAGTGTGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4669  
QY 4021 TGGAAATTCATAGCGGGATACAAATATTTAGAGGCTTGTGCACTCTGCTGCGGAACCCC 4080  
Db 4670 TGGAAATTCATAGCGGGATACAAATATTTAGAGGCTTGTGCACTCTGCTGCGGAACCCC 4729  
QY 4081 GCGATAGCATCACTGATGGCATTCAGAGCTCTATCACAGCGCGCTCACCCACCAACAT 4140  
Db 4730 GCGATAGCATCACTGATGGCATTCAGAGCTCTATCACAGCGCGCTCACCCACCAACAT 4789  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4200  
Db 4790 ACCCTCTGTTTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4849



Db 7010 CTGGAAGACTGAGACACCAATTGACACCAACCATATGCGCAAAAAATGAGTTTCTGCG 7069  
QY 6421 GTCCAAACAGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATTCGCCAGATTGGGG 6480  
Db 7070 GTCCAAACAGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATTCGCCAGATTGGGG 7129  
QY 6481 GTTCGTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTCCCTCAGGCCGCTG 6540  
Db 7130 GTTCGTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTCCCTCAGGCCGCTG 7189  
QY 6541 ATGGGCTCTTCAATCGGATTCGAATCTCTCTGACAGAGCGGGTTCGATTCCTGCTGAAT 6600  
Db 7190 ATGGGCTCTTCAATCGGATTCGAATCTCTCTGACAGAGCGGGTTCGATTCCTGCTGAAT 7249  
QY 6601 GCCTGAAAGCGAAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTCCCTCAGGCCGCTG 6660  
Db 7250 GCTTGGAAAGCGAAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTCCCTCAGGCCGCTG 7309  
QY 6661 ACGGTCACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTGTGACTTGGCC 6720  
Db 7310 ACGGTCACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTGTGACTTGGCC 7369  
QY 6721 CCGGAGCCAGACAGGCCATAGGTGCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 6780  
Db 7370 CCGGAGCCAGACAGGCCATAGGTGCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 7429  
QY 6781 ACTAATCTAAAGGCGAAGCTGGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Db 7430 ACTAATCTAAAGGCGAAGCTGGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 7489  
QY 6841 ACCAGCTCGGTAAATACCTCTCACATGTTACTTGAAGCGCGTCTCGGCTCTCGAGCTCGG 6900  
Db 7490 ACCAGCTCGGTAAATACCTCTCACATGTTACTTGAAGCGCGTCTCGGCTCTCGAGCTCGG 7549  
QY 6901 AAGCTCAGGACTGCAATGCTCTGATGCGGAGACCTTGTGCTTATCTGTAAGC 6960  
Db 7550 AAGCTCAGGACTGCAATGCTCTGATGCGGAGACCTTGTGCTTATCTGTAAGC 7609  
QY 6961 GCGGGGACCCAGAGAGAGCGGCGAGCTAGCGGCTTACGGAGGCTATGACTAGATAC 7020  
Db 7610 GCGGGGACCCAGAGAGAGCGGCGAGCTAGCGGCTTACGGAGGCTATGACTAGATAC 7669  
QY 7021 TCTGCCCCCTCGGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAAATCATGCG 7080  
Db 7670 TCTGCCCCCTCGGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAAATCATGCG 7729  
QY 7081 TCTTCCAAATGTGCTAGTGGCGCAGATGCTGTCGCAAAAGGGTGTACTATCTCACCCGT 7140  
Db 7730 TCTTCCAAATGTGCTAGTGGCGCAGATGCTGTCGCAAAAGGGTGTACTATCTCACCCGT 7789  
QY 7141 GACCCCAACCCCTTGGCGGGCTGCGTGGGAGCAGCTAGACACACTCCAGTCAAT 7200  
Db 7790 GACCCCAACCCCTTGGCGGGCTGCGTGGGAGCAGCTAGACACACTCCAGTCAAT 7849  
QY 7201 TCTTGGGTAGGCAACATCATGTATGCGGCCACCTTGTGGCAAGGATCATCTGATG 7260  
Db 7850 TCTTGGGTAGGCAACATCATGTATGCGGCCACCTTGTGGCAAGGATCATCTGATG 7909  
QY 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGAAAGAGCCCTAGATTGTGAG 7320  
Db 7910 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGAAAGAGCCCTAGATTGTGAG 7969  
QY 7321 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCATTCACGACTC 7380  
Db 7970 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCATTCACGACTC 8029  
QY 7381 CATGGCTTACGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
Db 8030 CACGGCTTACGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 8089  
QY 7441 TCATGCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAT 7500  
Db 8090 TCATGCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAT 8149

QY 7501 GTCCGCGCTAGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 8150 GTCCGCGCTAGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 8209  
QY 7561 AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCAGTTGGAT 7620  
Db 8210 AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCAGTTGGAT 8269  
QY 7621 TTATCCAGCTGGTTCTGTTGTGTTTACAGCGGGGAGACATATATCACAGCCCTGTCTGT 7680  
Db 8270 TTATCCAGCTGGTTCTGTTGTGTTTACAGCGGGGAGACATATATCACAGCCCTGTCTGT 8329  
QY 7681 GCCCGACCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7740  
Db 8330 GCCCGACCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8389  
QY 7741 CTACTCCCCAACCCGATGAACGGGACCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 7800  
Db 8390 CTACTCCCCAACCCGATGAACGGGAGCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 8449  
QY 7801 TTTCCC-----TTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7850  
Db 8450 TTTCCCCTTTTTTCCCTTT 8509  
QY 7851 TTTCCCCTTTTTTCCCTTT 7910  
Db 8510 CCCCCCTTTTTTCCCTTT 8569  
QY 7911 TAGTCACGGCTAGCTGTGAAAGTCCGTGAGCGCTTGAAGTCCGTGAGAGAGTCTGATCTG 7970  
Db 8570 TAGTCACGGCTAGCTGTGAAAGTCCGTGAGCGCTTGAAGTCCGTGAGAGAGTCTGATCTG 8629  
QY 7971 GCCTCTCTGCAGATCAAGT 7989  
Db 8630 GCCTCTCTGCAGATCAAGT 8648

RESULT 15  
US-10-029-907-4  
; Sequence 4, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8643  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-4

Query Match 90.4%; Score 7221; DB 4; Length 8643;  
Best Local Similarity 91.9%; Pred. No. 0;  
Matches 7952; Conservative 0; Mismatches 25; Indels 676; Gaps 4;

QY 2 CCAGCCCCCATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTGT 61  
Db 2 CCAGCCCCCATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTGT 61  
QY 62 CTTTACGCAAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGAGCTCCAGGACC 121  
Db 62 CTTTACGCAAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGAGCTCCAGGACC 121

122 QY CCCCCTCCGGAGACCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAGG 181  
122 Db CCCCCTCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAGG 181  
182 QY ACAGCCGGGTCCCTTTCTTGATCAACCCGCTCAATGCTTGGGATTTGGGCGGTGCCCCCG 241  
182 Db ACAGCCGGGTCCCTTTCTTGATCAACCCGCTCAATGCTTGGGATTTGGGCGGTGCCCCCG 241  
242 QY CGAGACTGTAGCCGAGTGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGTAGAGG 301  
242 Db CGAGACTGTAGCCGAGTGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGTAGAGG 301  
302 QY TGCCTTCGAGTCCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 361  
302 Db TGCCTTCGAGTCCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 361  
362 QY TCAAGAAAAACCAAGGGCGCCCATGATTGAACAAGATGGAATGACGACGAGTTCTCC 421  
362 Db TCAAGAAAAACCAAGGGCGCCCATGATTGAACAAGATGGAATGACGACGAGTTCTCC 421  
422 QY GGCCTGTTGGTGGAGGCTATTCCGCTATGATGCTGGGCAACACAGACATCGGCTGCTC 481  
422 Db GGCCTGTTGGTGGAGGCTATTCCGCTATGATGCTGGGCAACACAGACATCGGCTGCTC 481  
482 QY TGAATCCGCGTCTCCGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGA 541  
482 Db TGAATCCGCGTCTCCGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGA 541  
542 QY CTTGTCGGGTGCTCCTGAATGAATGACGAGGAGGCGCGCTATCGTGGTGGCCAC 601  
542 Db CTTGTCGGGTGCTCCTGAATGAATGACGAGGAGGCGCGCTATCGTGGTGGCCAC 601  
602 QY GACGGCGTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGGAAGGACTGGCT 661  
602 Db GACGGCGTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGGAAGGACTGGCT 661  
662 QY GCTATTGGGGAAGTCCCGGGGAGGATCTCTGTGATCTACCTGCTCCTCCCGAGAA 721  
662 Db GCTATTGGGGAAGTCCCGGGGAGGATCTCTGTGATCTACCTGCTCCTCCCGAGAA 721  
722 QY AGTATCCATCATGCTGATGATGCGGCGGCTGCTACGCTTGTATCGGCTACCTGCC 781  
722 Db AGTATCCATCATGCTGATGATGCGGCGGCTGCTACGCTTGTATCGGCTACCTGCC 781  
782 QY ATTGACACCAAGCGAAACATCGATCGAGCGAGCACTGCTGGAATGGAAGCGGTCT 841  
782 Db ATTGACACCAAGCGAAACATCGATCGAGCGAGCACTGCTGGAATGGAAGCGGTCT 841  
842 QY TGTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTGCG 901  
842 Db TGTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTGCG 901  
902 QY CAGGCTCAAGGCGCGATGCGGACGCGAGGATCTGCTGTCGACCATGCGGATGCTG 961  
902 Db CAGGCTCAAGGCGCGATGCGGACGCGAGGATCTGCTGTCGACCATGCGGATGCTG 961  
962 QY CTTGCCGAATATCATGTTGAAAATGCGCGCTTTTCTGGAATTCATCGACTGTGCGCGGCT 1021  
962 Db CTTGCCGAATATCATGTTGAAAATGCGCGCTTTTCTGGAATTCATCGACTGTGCGCGGCT 1021  
1022 QY GGTGTGGGACCGCTATCAGACATAGCGTTGGCTACCGCTGATATGCTGAAGAGCT 1081  
1022 Db GGTGTGGGACCGCTATCAGACATAGCGTTGGCTACCGCTGATATGCTGAAGAGCT 1081  
1082 QY TGGCGGGAATGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCGCA 1141  
1082 Db TGGCGGGAATGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCGCA 1141  
1142 QY GCGCATCGCTTCTATCGCTTCTTGAAGGAGTCTTCTGAGT-----TAAAC 1189  
1142 Db GCGCATCGCTTCTATCGCTTCTTGAAGGAGTCTTCTGAGT-----TAAAC 1201

1190 QY AGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTA 1249  
1202 Db AGACCAACAGGTTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCTA 1250  
1250 QY AGTTACTGGCGAAGCGCTTGGGAATAAGGCGGTGTGCGTTTGTCTATATATTTT 1309  
1251 Db AGTTACTGGCGAAGCGCTTGGGAATAAGGCGGTGTGCGTTTGTCTATATATTTT 1310  
1310 QY CCACCATATTTGCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTCTTGA 1369  
1311 Db CCACCATATTTGCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTCTTGA 1370  
1370 QY CGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATTCG 1429  
1371 Db CGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATTCG 1430  
1430 QY TGAAGGAAGAGTTCTCTGGAAGCTTCTTGAAGCAAAACAAGCTGTAGGACCTTT 1489  
1431 Db TGAAGGAAGAGTTCTCTGGAAGCTTCTTGAAGCAAAACAAGCTGTAGGACCTTT 1490  
1490 QY GCAGGAGCGGAACCCCTGCGACAGGTGCTCTGCGGCCAAAGCCACGCTGAT 1549  
1491 Db GCAGGAGCGGAACCCCTGCGACAGGTGCTCTGCGGCCAAAGCCACGCTGAT 1550  
1550 QY AAGATACCTGCAAGGCGGCAACCCCTGCGACAGGTGCTGAGTTGATAGTTGCG 1609  
1551 Db AAGATACCTGCAAGGCGGCAACCCCTGCGACAGGTGCTGAGTTGATAGTTGCG 1610  
1610 QY AAGAGTCAAAATGGCTCTCTGAGCGTATTCAAGGGGCTCAAGGATGCCAGAGG 1669  
1611 Db AAGAGTCAAAATGGCTCTCTGAGCGTATTCAAGGGGCTCAAGGATGCCAGAGG 1670  
1670 QY TACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGT 1729  
1671 Db TACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGT 1730  
1730 QY CGAGGTTTAAAAACGTTAGGCGGCGGAAACCAAGGGGACGTTGTTCTTTGAAAAAC 1789  
1731 Db CGAGGTTTAAAAACGTTAGGCGGCGGAAACCAAGGGGACGTTGTTCTTTGAAAAAC 1790  
1790 QY ACGATTAATACCATG----- 1803  
1791 Db ACGATTAATACCATGAGACCGGAGATGCGAGCATCTGTCGAGGCGCGGTTTCTGAGTCT 1850  
1804 QY ----- 1803  
1851 Db TGAATCTTTGACCTTGTACCGCATTAAGCTGTCTCGCTAGGCTCATATGTTGTT 1910  
1804 QY ----- 1803  
1911 Db TACAATATTTATCACCAGGCGGAGGACACATTGCAAGTGTGGATCCCCCCTCAACG 1970  
1804 QY ----- 1803  
1971 Db TTGCGGGGCGCGGATGCGGCTCATCTCTACGTCGCGATCCACCCAGAGCTAATCT 2030  
1804 QY ----- 1803  
2031 Db TTACCATCACCAAAATCTTGTCTGCCATACTCGTCCACTCATGTTGCTCCAGGCTGTA 2090  
1804 QY ----- 1803  
2091 Db TAACCAAGTGCCTACTTCTGTCGCGCACACGCGGCTCATTCGTGATGATGCTGTTGTC 2150  
1804 QY ----- 1803  
2151 Db GGAAGTTGCTGGGGTCAATTATGTCCAAATGCTCTCATGAAGTTGGCGGCTGACAG 2210  
1804 QY ----- 1803  
2211 Db GTAGTACGTTTATGACCATCTCACCCACTGCGGGACTGGGCCACGCGGCGCTACAG 2270  
1804 QY ----- 1803

Db 2271 ACCTTGGCTGGCAGTTGAGCCCGTGGTCTTCTCTGATATGAGACCAAGGTTATCACT 2330  
QY 1804 ----- 1803  
Db 2331 GGGGGGACAGACACCGCGCGTGTGGGGACATCATCTTGGGCCCTGCCCGTCTCCGCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGGAGGAGATACATCTGGGACCGGACAGAGCCCTTGAAGGGCAGGGTGGGACTCC 2450  
QY 1804 --GGCCCTATTACGCCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACTA 1861  
Db 2451 TCGCGCCTATTACGCCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACTA 2510  
QY 1862 GCCTCAGAGCGGGACAGGAACAGAGTCGAGGGGGAGGTCCAAATGTCTCCACCGAA 1921  
Db 2511 GCCTCAGAGCGGGACAGGAACAGAGTCGAGGGGGAGGTCCAAATGTCTCCACCGAA 2570  
QY 1922 CACATCTTCTCGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGGCG 1981  
Db 2571 CACATCTTCTCGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGGCG 2630  
QY 1982 GCTCAAGACCCCTTGC CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACC 2041  
Db 2631 GCTCAAGACCCCTTGC CGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACC 2690  
QY 2042 AGGACCTCGTGGCTGGCAAGGCCGCCCGGGGCGCTTCTTGACACCATGCACTGCG 2101  
Db 2691 AGGACCTCGTGGCTGGCAAGGCCGCCCGGGGCGCTTCTTGACACCATGCACTGCG 2750  
QY 2102 GCAGCTCGGACTTACTTGTGTCAGAGGCATGCCGATGTCATTCGGTGGCGCGCGG 2161  
Db 2751 GCAGCTCGGACTTACTTGTGTCAGAGGCATGCCGATGTCATTCGGTGGCGCGCGG 2810  
QY 2162 GCAGCAGAGGGGAGCCTACTCTCCCGAGCCCGTCTCCTACTTGAAGGCTCTTCGG 2221  
Db 2811 GCAGCAGAGGGGAGCCTACTCTCCCGAGCCCGTCTCCTACTTGAAGGCTCTTCGG 2870  
QY 2222 GCGTCCACTGCTCTGCCCCCTGGGGCAGCTGTGGGCATCTTTCGGGCTGCGGTGTGCA 2281  
Db 2871 GCGTCCACTGCTCTGCCCCCTGGGGCAGCTGTGGGCATCTTTCGGGCTGCGGTGTGCA 2930  
QY 2282 CCCGAGGGTTGCGAAGCGGTGACCTTTGACCGTGCAGTCTATGGAACCACTATGC 2341  
Db 2931 CCCGAGGGTTGCGAAGCGGTGACCTTTGACCGTGCAGTCTATGGAACCACTATGC 2990  
QY 2342 GGTCCCGGCTTTCAGGACAACTCGTCCCTCCGGCGGTACCGCAGACATTCAGGTGG 2401  
Db 2991 GGTCCCGGCTTTCAGGACAACTCGTCCCTCCGGCGGTACCGCAGACATTCAGGTGG 3050  
QY 2402 CCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCGCGCTCGGTATGAG 2461  
Db 3051 CCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCGCGCTCGGTATGAG 3110  
QY 2462 CCCAAGGATATAAGTGCTTCTGCTGAACCGTCCGTCGCGCCACCTAGGTTTCGGG 2521  
Db 3111 CCCAAGGATATAAGTGCTTCTGTAACCCGTCCGTCCCGCCACCTAGGTTTCGGG 3170  
QY 2522 CGTATATGCTAAGGCAATGGTATCGACCTTAACATCAGAACCGGGGTGAAGCACTCA 2581  
Db 3171 CGTATATGCTAAGGCAATGGTATCGACCTTAACATCAGAACCGGGGTGAAGCACTCA 3230  
QY 2582 CCAGGTTGCCCTCATCAGTACTCCACCTATGGCAAGTTTCTTCCGACCGTGGTGTGCT 2641  
Db 3231 CCAGGTTGCCCTCATCAGTACTCCACCTATGGCAAGTTTCTTCCGACCGTGGTGTGCT 3290  
QY 2642 CTGGGGCGCTATGACATCATATATGTGATGAGTGCACCTCACTGACTTCGACCACTA 2701  
Db 3291 CTGGGGCGCTATGACATCATATATGTGATGAGTGCACCTCACTGACTTCGACCACTA 3350  
QY 2702 TCCTGGGCATCGGCACAGTCTCGGACCAAGCGGAGAGCGGCTGGAGCGGCTCGTCTGTC 2761

Db 3351 TCCTGGGCATCGGCACAGTCTCTGGACCAAGCGGACAGCGCTGGAGCGGACTCGTCTGTC 3410  
QY 2762 TCGCCACCGCTACCGCTCCGGGATCGGTCAACGTCGCCATCCAAACATCGAGGAGTGG 2821  
Db 3411 TCGCCACCGCTACCGCTCCGGGATCGGTCAACGTCGCCATCCAAACATCGAGGAGTGG 3470  
QY 2822 CTCCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCA 2881  
Db 3471 CTCCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCA 3530  
QY 2882 AGGGGGGAGGACCTCTATTTCTGCCATTCACAAAGAAATGTGATGAGTTCGCGCA 2941  
Db 3531 AGGGGGGAGGACCTCTATTTCTGCCATTCACAAAGAAATGTGATGAGTTCGCGCA 3590  
QY 2942 AGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATCCGTCA 3001  
Db 3591 AGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATCCGTCA 3650  
QY 3002 TACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTCTAATGACGGGCTTTACG 3061  
Db 3651 TACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTCTAATGACGGGCTTTACG 3710  
QY 3062 GCGATTTGCACTCAGTGTGATGACCTGCAATACATGTGTCAACGAGACGTTCGACTTCAGCC 3121  
Db 3711 GCGATTTGCACTCAGTGTGATGACCTGCAATACATGTGTCAACGAGACGTTCGACTTCAGCC 3770  
QY 3122 TGGACCCGACCTTACCATTGAGACGACGACCGTGTGCCAAGAGCGGCTGTACGCTCGC 3181  
Db 3771 TGGACCCGACCTTACCATTGAGACGACGACCGTGTGCCAAGAGCGGCTGTACGCTCGC 3830  
QY 3182 AGCGCGAGGACGAGCTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGATCCAGAG 3241  
Db 3831 AGCGCGAGGACGAGCTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGATCCAGAG 3890  
QY 3242 AACGGCCCTCGGGCATGTTTCGATTCCTCGTGTGTGTGAGTGTATCACCGGCTGTG 3301  
Db 3891 AACGGCCCTCGGGCATGTTTCGATTCCTCGTGTGTGTGAGTGTATCACCGGCTGTG 3950  
QY 3302 CTTGGTAGAGCTCAACCGCCCGAGACCTCAGTTAGGTTCGGGCTTACCTAAACACAC 3361  
Db 3951 CTTGGTAGAGCTCAACCGCCCGAGACCTCAGTTAGGTTCGGGCTTACCTAAACACAC 4010  
QY 3362 CAGGTTTCCCGCTGTCAGGACCATCTGAGATTTCTGGGAGAGGTCTTTACAGGCTCA 3421  
Db 4011 CAGGTTTCCCGCTGTCAGGACCATCTGAGATTTCTGGGAGAGGTCTTTACAGGCTCA 4070  
QY 3422 CCCACATAGAGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC 3481  
Db 4071 CCCACATAGAGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC 4130  
QY 3482 TGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGGGACC 3541  
Db 4131 TGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGGGACC 4190  
QY 3542 AAATGTGGAGTGTCTCATACGGCTAAAGCCTACGCTGCACGGGCCAACGCCCTGTGT 3601  
Db 4191 AAATGTGGAGTGTCTCATACGGCTAAAGCCTACGCTGCACGGGCCAACGCCCTGTGT 4250  
QY 3602 ATAGGCTGGGAGCCGTTCAAAACGAGGTCTACACACACACCCCATCAACCAATACATCA 3661  
Db 4251 ATAGGCTGGGAGCCGTTCAAAACGAGGTCTACACACACACCCCATCAACCAATACATCA 4310  
QY 3662 TGGCATGCATGTGGCTGACCTGAGGTGTGCTGACGACCTGCGGTCTGTTAGGCGGAG 3721  
Db 4311 TGGCATGCATGTGGCTGACCTGAGGTGTGCTGACGACCTGCGGTCTGTTAGGCGGAG 4370  
QY 3722 TCCTAGCAGCTCTGGCCGCGTATTCCTGACACACAGCAGCGGTGCTCATTTGTGGGACGA 3781  
Db 4371 TCCTAGCAGCTCTGGCCGCGTATTCCTGACACACAGCAGCGGTGCTCATTTGTGGGACGA 4430  
QY 3782 TCATTTGTCCGAAAGCGCGCATCATTCGCCACAGGAGTCCCTTTACGGGAGTTTCG 3841  
Db 4431 TCATTTGTCCGAAAGCGCGCATCATTCGCCACAGGAGTCCCTTTACGGGAGTTTCG 4490









**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:37:05 ; Search time 3599 Seconds  
(without alignments)  
12238.821 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989  
Sequence: 1 gccagccccgattgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA: \*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_PCT\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PCT\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7987.4	100.0	7989	17	US-10-639-150-1
2	7987.4	100.0	7992	13	US-10-005-469-1
3	7987.4	100.0	7992	16	US-10-434-842-1
4	7987.4	100.0	10690	15	US-10-125-940-1
5	7987.4	100.0	10690	14	US-10-125-920-1
6	7987.4	100.0	10690	16	US-10-467-000-3
7	7985.8	100.0	7992	13	US-10-005-469-6
8	7985.8	100.0	7992	16	US-10-434-842-6
9	7982.6	99.9	7992	13	US-10-005-469-2
10	7982.6	99.9	7992	13	US-10-005-469-5
11	7982.6	99.9	7992	16	US-10-434-842-2
12	7982.6	99.9	7992	16	US-10-434-842-5

13	7982.6	99.9	7992	16	US-10-434-842-17	Sequence 17, Appl
14	7981	99.9	7992	13	US-10-005-469-4	Sequence 4, Appl
15	7981	99.9	7992	16	US-10-434-842-4	Sequence 4, Appl
16	7981	99.9	7992	16	US-10-434-842-15	Sequence 15, Appl
17	7971.4	99.8	7989	13	US-10-434-842-16	Sequence 16, Appl
18	7971.2	99.8	7995	13	US-10-005-469-3	Sequence 3, Appl
19	7971.2	99.8	7995	16	US-10-434-842-3	Sequence 3, Appl
20	7610.4	95.3	12305	18	US-10-422-323A-2	Sequence 2, Appl
21	7607.2	95.2	12315	18	US-10-422-323A-1	Sequence 1, Appl
22	7272.2	91.0	8639	13	US-10-029-907-1	Sequence 1, Appl
23	7272.2	91.0	8639	15	US-10-309-561-1	Sequence 1, Appl
24	7272.2	91.0	8639	17	US-10-789-355-1	Sequence 1, Appl
25	7272.2	91.0	8639	18	US-10-686-835-1	Sequence 1, Appl
26	7271.2	91.0	8638	13	US-10-029-907-24	Sequence 24, Appl
27	7271.2	91.0	8638	15	US-10-309-561-24	Sequence 24, Appl
28	7271.2	91.0	8638	17	US-10-789-355-24	Sequence 24, Appl
29	7271.2	91.0	8638	18	US-10-686-835-24	Sequence 24, Appl
30	7248.2	90.7	8638	13	US-10-029-907-7	Sequence 7, Appl
31	7248.2	90.7	8638	15	US-10-309-561-7	Sequence 7, Appl
32	7248.2	90.7	8638	17	US-10-789-355-7	Sequence 7, Appl
33	7248.2	90.7	8638	18	US-10-686-835-7	Sequence 7, Appl
34	7247.2	90.7	8638	13	US-10-029-907-25	Sequence 25, Appl
35	7247.2	90.7	8638	15	US-10-309-561-25	Sequence 25, Appl
36	7247.2	90.7	8638	17	US-10-789-355-25	Sequence 25, Appl
37	7247.2	90.7	8638	18	US-10-686-835-25	Sequence 25, Appl
38	7245.2	90.7	8642	13	US-10-029-907-2	Sequence 2, Appl
39	7245.2	90.7	8642	15	US-10-309-561-2	Sequence 2, Appl
40	7245.2	90.7	8642	17	US-10-789-355-2	Sequence 2, Appl
41	7245.2	90.7	8642	18	US-10-686-835-2	Sequence 2, Appl
42	7244	90.7	8638	13	US-10-029-907-6	Sequence 6, Appl
43	7244	90.7	8638	15	US-10-309-561-6	Sequence 6, Appl
44	7244	90.7	8638	17	US-10-789-355-6	Sequence 6, Appl
45	7244	90.7	8638	18	US-10-686-835-6	Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-10-639-150-1  
; Sequence 1, Application US/10639150  
; Publication No. US20040121975A1  
; GENERAL INFORMATION:  
; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY  
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS  
; FILE REFERENCE: D0224 NP  
; CURRENT APPLICATION NUMBER: US/10/639,150  
; CURRENT FILING DATE: 2003-08-12  
; PRIOR APPLICATION NUMBER: US 60/402,661  
; PRIOR FILING DATE: 2002-08-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1  
; LENGTH: 7989  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: HCV Replicon  
US-10-639-150-1

Query Match	100.0%	Score	7987.4;	DB	17;	Length	7989;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	7988;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
Qy	1	GCACGCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG	60				
Db	1	GCACGCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG	60				
Qy	61	TCTTCACGAGAAACGGTCTAGCCATGCGTGTAGTATGAGTGTGCGTGCAGCCTCCAGAC	120				
Db	61	TCTTCACGAGAAACGGTCTAGCCATGCGTGTAGTATGAGTGTGCGTGCAGCCTCCAGAC	120				
Qy	121	CCCCCTCCCGGAGAGCCATAGTGTGCGGAAACCGGTGAGTACACCGGAATTCGAC	180				

121 CCCCCCTCCCGGAGGACATAGTGGTCTCGGGAACGGGTAGTACACCGGAATTGGCAG 180  
 181 GAGACCGGGTCTTCTTGGATCAACCCCGCTCAATGCTTGGAGATTGGGCGTGGCCCC 240  
 181 GACGACCGGGTCTTCTTGGATCAACCCCGCTCAATGCTTGGAGATTGGGCGTGGCCCC 240  
 241 GCGAGACTGCTAGCGAGTAGTCTTGGGTGCGCAAGAGCCCTTGTGTACTGCTCCGTATAG 300  
 241 GCGAGACTGCTAGCGAGTAGTCTTGGGTGCGCAAGAGCCCTTGTGTACTGCTCCGTATAG 300  
 301 GTCTTTGCGAGTGCCTCCGCGGAGTCTCTGTAGACCGGTGCACCATGAGCACGAATCTTAAC 360  
 301 GTCTTTGCGAGTGCCTCCGCGGAGTCTCTGTAGACCGGTGCACCATGAGCACGAATCTTAAC 360  
 361 CTCAAGAAACCAAGAGGCGGCCCATGATTTGAAACAAGATGATTTGCAACGAGTCTTC 420  
 361 CTCAAGAAACCAAGAGGCGGCCCATGATTTGAAACAAGATGATTTGCAACGAGTCTTC 420  
 421 CGCCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCGGCGGCGGCTATCGTGGCTGCT 480  
 421 CGCCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCGGCGGCGGCTATCGTGGCTGCT 480  
 481 CTGATCCCGCGGTCTCCGCTGTACGCGAGGCGGCGGCGGCTCTTTTGTCAAGACCG 540  
 481 CTGATCCCGCGGTCTCCGCTGTACGCGAGGCGGCGGCGGCTCTTTTGTCAAGACCG 540  
 541 ACTGTCCGGTCCCTGATGAATGAACTGAGGAGGCGGCGGCGGCTATCGTGGCGGCA 600  
 541 ACTGTCCGGTCCCTGATGAATGAACTGAGGAGGCGGCGGCGGCTATCGTGGCGGCA 600  
 601 CGACGGCGGTCTCTCCGAGCTGTCTGACGCTGTCTGATCTCATCTTGTCTTCCGCGAGA 720  
 601 CGACGGCGGTCTCTCCGAGCTGTCTGACGCTGTCTGATCTCATCTTGTCTTCCGCGAGA 720  
 661 TGCTATTGGCGAAGTGCGCGGCGGAGTCTCTGTCTCATCTTGTCTTCCGCGAGA 720  
 721 AGTATCCATATGCGTGTATGCAATGCGGGGCTGATACGCTGTGATTCGGCTACTGCG 780  
 721 AGTATCCATATGCGTGTATGCAATGCGGGGCTGATACGCTGTGATTCGGCTACTGCG 780  
 781 CATTCGACCAACGAGGAACTCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTC 840  
 781 CATTCGACCAACGAGGAACTCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTC 840  
 841 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAATCTTCG 900  
 841 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAATCTTCG 900  
 901 CCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCGCT 960  
 901 CCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCGCT 960  
 961 GCTTCCGGAATCATGTGTGAAATGCGGCTTTCTGATTCATCGACTGTGGCGGCG 1020  
 961 GCTTCCGGAATCATGTGTGAAATGCGGCTTTCTGATTCATCGACTGTGGCGGCG 1020  
 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATTGCTGAAGAGC 1080  
 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATTGCTGAAGAGC 1080  
 1081 TTGGGGGCAATGGGCTGACCGCTCTCTGCTGTATACGGTATCGCGCTCCGGATTCG 1140  
 1081 TTGGGGGCAATGGGCTGACCGCTCTCTGCTGTATACGGTATCGCGCTCCGGATTCG 1140  
 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGATTTCTGAGTTTAAACAGACCAACAG 1200  
 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGATTTCTGAGTTTAAACAGACCAACAG 1200  
 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCGCTTAACGTTACTGGC 1260

1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTAACGTTACTGGC 1260  
 1261 CGAAGCGCTTGGAAATGAGCGCGGTGTGGTTTGTCTATATGTTATTTTCCACATATTG 1320  
 1261 CGAAGCGCTTGGAAATGAGCGCGGTGTGGTTTGTCTATATGTTATTTTCCACATATTG 1320  
 1321 CGGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTGAGGAGCATTCCT 1380  
 1321 CGGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTGAGGAGCATTCCT 1380  
 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCTGTGAAGGAAGCA 1440  
 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCTGTGAAGGAAGCA 1440  
 1441 GTTCTCTGGAAGTCTTTGGAAGCAAAACAAGTCTGTAGCGACCCCTTTGAGGCGCGG 1500  
 1441 GTTCTCTGGAAGTCTTTGGAAGCAAAACAAGTCTGTAGCGACCCCTTTGAGGCGCGG 1500  
 1501 AACCCCGCCACCTGGCGACAGGTGCTCTCGGCGCAAAAGCCACGTGTATAAGATACACCT 1560  
 1501 AACCCCGCCACCTGGCGACAGGTGCTCTCGGCGCAAAAGCCACGTGTATAAGATACACCT 1560  
 1561 GCAAGGCGGCAACACCCAGTGCACCGTTGTGAGTGGATAGTTGTGAAAAGAGTCAAA 1620  
 1561 GCAAGGCGGCAACACCCAGTGCACCGTTGTGAGTGGATAGTTGTGAAAAGAGTCAAA 1620  
 1621 TGGCTCTCTCAAGCGTATTCAAGGGGCTGAGAGTGCAGAGAGTACCCCATTTGT 1680  
 1621 TGGCTCTCTCAAGCGTATTCAAGGGGCTGAGAGTGCAGAGAGTACCCCATTTGT 1680  
 1681 ATGGGATCTGATCTGGGCGCTCGTGACATGCTTTACATGCTTTAGTTCGAGTTAAAA 1740  
 1681 ATGGGATCTGATCTGGGCGCTCGTGACATGCTTTACATGCTTTAGTTCGAGTTAAAA 1740  
 1741 AACGTCTAGGCGCCCGCAACACCGGAGCTGTGTTTCTTTGAAAAACACGATATAACC 1800  
 1741 AACGTCTAGGCGCCCGCAACACCGGAGCTGTGTTTCTTTGAAAAACACGATATAACC 1800  
 1801 ATGGCGCTTATTAGCGCTTACTCCCAACAGACGCGAGGCGCTTCTGGCTGCATCATCACT 1860  
 1801 ATGGCGCTTATTAGCGCTTACTCCCAACAGACGCGAGGCGCTTCTGGCTGCATCATCACT 1860  
 1861 AGCTCACAGCGCGGACAGGACACAGGTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
 1861 AGCTCACAGCGCGGACAGGACACAGGTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
 1921 ACACATCTTCTTCTGGCGACCTGCGTCAATGCGGTGTGTTGGACTGCTATCATGTGGCC 1980  
 1921 ACACATCTTCTTCTGGCGACCTGCGTCAATGCGGTGTGTTGGACTGCTATCATGTGGCC 1980  
 1981 GGCTCAAGACCTTTCGCGGCGCAAGGGCGCAATCACCAATGTACCAATGTGGAC 2040  
 1981 GGCTCAAGACCTTTCGCGGCGCAAGGGCGCAATCACCAATGTACCAATGTGGAC 2040  
 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGAACCATGCACCTGC 2100  
 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGAACCATGCACCTGC 2100  
 2101 GGCAGCTCGGACCTTTTACTTGGTTCAGAGGATGCGGATGTCAATTCGGGTGCGCGCGG 2160  
 2101 GGCAGCTCGGACCTTTTACTTGGTTCAGAGGATGCGGATGTCAATTCGGGTGCGCGCGG 2160  
 2161 GCGGACAGGAGGGAGCGCTACTCTCCCGCGAGCGCGCTCTTCTTGAAGGCTCTTTCG 2220  
 2161 GCGGACAGGAGGGAGCGCTACTCTCCCGCGAGCGCGCTCTTCTTGAAGGCTCTTTCG 2220  
 2221 GCGGCTCCACTGCTCTGCGGCTCGGGGACCGTGTGGGATCTTTTGAAGGCTCTTTCG 2280  
 2221 GCGGCTCCACTGCTCTGCGGCTCGGGGACCGTGTGGGATCTTTTGAAGGCTCTTTCG 2280  
 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGTCTGAGTCTATGAAAACCACTATG 2340  
 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGTCTGAGTCTATGAAAACCACTATG 2340

QY 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGCGCTAACCGCAGACATTCAGGTG 2400  
Db  
QY 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGCGCTAACCGCAGACATTCAGGTG 2400  
Db  
QY 2401 GCCCATCTACACGCCCTTACTGGTAGGGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 2460  
Db  
QY 2401 GCCCATCTACACGCCCTTACTGGTAGGGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 2460  
QY 2461 GCCCAAGGGTATAAGGTGCTTGTCTCTGAA CCGCTCGCTCGCGCA CCGTAGGTTTCGGG 2520  
Db  
QY 2521 GCGTATATGCTAAGGCACATGGTATCGACCTTAACATCAACATCAAGACCGGGTAAAGACCATC 2580  
Db  
QY 2581 ACCACGGGTGCCCCATCAGTACTCCACCTATGGCAAGTTTCTTGCGGACGGTGTTC 2640  
Db  
QY 2641 TCTGGGGCGGCTATGATCATATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db  
QY 2701 ATCTGGGCATCGGCACAGTCTCGGATCGGTCAACCGTGCACATCAAAACATCGAGAGGTG 2760  
Db  
QY 2761 CTGCCACCGCTACGCTCCGGATCGGTCAACCGTGCACATCAAAACATCGAGAGGTG 2820  
QY 2821 GCTCTGTCAGCACTGAGAAATCCCTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db  
QY 2881 AAGGGGGAGGACCTCATTTCTGCCATTCGAAGAAATGTGATGAGCTCGCGCG 2940  
Db  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTACCGGGGCTTGATGTATCGCTC 3000  
Db  
QY 3001 ATACCAACTAGCGGAGAGCTCAATGTGTAGCAACCGAGCGCTTAATGACGGGTTTACC 3060  
Db  
QY 3061 GGGCATTTGCACTCAGTGATCGACTGCAATACATGTGTCAACCGAGACAGTCACTTACG 3120  
Db  
QY 3121 CTGACCCGACCTTCACTTGAACAGCAGACCGTGCACAAAGACGGGTGTCACTGTCG 3180  
Db  
QY 3181 CAGCGGAGGACGACTGTGTAGGGGAGGATGGGCATTTACAGGTTGTGACTCCAGGA 3240  
Db  
QY 3241 GAACGGCCCTCGGCATGTTGATTCCTCGTTCTGTGCGAGTCTATGACCGGGCTGT 3300  
Db  
QY 3301 GCTTTGGTACGAGCTCAGCCCGCGCAGACCTCAGTTAGTTTGGGGCTTACTAAACACA 3360  
Db  
QY 3361 CCAGGTTGCCGTCTGCGAGGACCATCTGAGTTCCTGGAGAGCGTCTTTACAGGCTC 3420  
Db  
QY 3421 ACCACATAGACGCCCACTTTCTTCTCCAGACTAAGCAGGACGAGACAACCTTCCCTAC 3480  
Db  
QY 3421 ACCACATAGACGCCCACTTTCTTCTCCAGACTAAGCAGGACGAGACAACCTTCCCTAC 3480  
Db  
QY 3481 CTGGTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Db  
QY 3481 CTGGTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
QY 3541 CAATGTGGAGTGTCTCATACGGCTAAAGCTTAACGCTTAACGCTGACGCGGCAACCGCCCTGCTG 3600  
Db  
QY 3601 TATAGGCTGGAGCCGTTCAAAACGAGGTTACTACCAACACACCCCAATAACCAATACATC 3660  
Db  
QY 3661 ATGSCATGTCATGTGCGGTGACCTCGAGGTGCTCACAGGACCTCGGTGCTGTAGGCGGA 3720  
Db  
QY 3721 GTCCTAGAGCTCTGGCGCGTATGCTGACAAACAGGAGCTGTGCTATTTGTGGGAGG 3780  
Db  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTCGCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTCGCGACAGGGAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGCCTCACACTCCCTTACATCGAACAGGGAATCAGCTGCGC 3900  
Db  
QY 3901 GAACAAATTCAAACAGAGGCAATCGGGTGTGTCGCAACAGCACCAGCAGGAGGCT 3960  
Db  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTCGGGGAGCATATG 4020  
Db  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTCGGGGAGCATATG 4020  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCTCTGCTTGGCAACCCC 4080  
Db  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCTCTGCTTGGCAACCCC 4080  
QY 4081 GCGATAGCATCATGATGGCATTCACAGCCTTATACACAGCCTGATACACCCCAACAT 4140  
Db  
QY 4081 GCGATAGCATCATGATGGCATTCACAGCCTTATACACAGCCTGATACACCCCAACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTCGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
Db  
QY 4141 ACCCTCTGTTTAAACATCTCGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTCTTTTCTGAGCGCGCATCGCTGGAGCGCTGTGTGGCAGCATAGGCTTTGGG 4260  
Db  
QY 4261 AAGTGTCTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGGCTGCTGGCC 4320  
Db  
QY 4261 AAGTGTCTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGGCTGCTGGCC 4320  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT 4380  
Db  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGCGCCCTAGTCTGCGGGTGTGTGGCAGCATGCTGCGTGGCAC 4440  
Db  
QY 4441 GTGGGCCAGGAGGAGGGGCTGTGCACTGATGAA CCGGCTGATAGCGTTTCCGTTCCGGG 4500  
Db  
QY 4441 GTGGGCCAGGAGGAGGGGCTGTGCACTGATGAA CCGGCTGATAGCGTTTCCGTTCCGGG 4500  
QY 4501 GGTAAACACGCTCTCCCCCAGCATATGTGCTGAGAGCGGCTGCGACGCTGCTCACT 4560

QY 3421 ACCACATAGACGCCCACTTTCTTCTCCAGACTAAGCAGGACGAGACAACCTTCCCTAC 3480  
Db  
QY 3421 ACCACATAGACGCCCACTTTCTTCTCCAGACTAAGCAGGACGAGACAACCTTCCCTAC 3480  
Db  
QY 3481 CTGGTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Db  
QY 3481 CTGGTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
QY 3541 CAATGTGGAGTGTCTCATACGGCTAAAGCTTAACGCTTAACGCTGACGCGGCAACCGCCCTGCTG 3600  
Db  
QY 3601 TATAGGCTGGAGCCGTTCAAAACGAGGTTACTACCAACACACCCCAATAACCAATACATC 3660  
Db  
QY 3661 ATGSCATGTCATGTGCGGTGACCTCGAGGTGCTCACAGGACCTCGGTGCTGTAGGCGGA 3720  
Db  
QY 3721 GTCCTAGAGCTCTGGCGCGTATGCTGACAAACAGGAGCTGTGCTATTTGTGGGAGG 3780  
Db  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTCGCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTCGCGACAGGGAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGCCTCACACTCCCTTACATCGAACAGGGAATCAGCTGCGC 3900  
Db  
QY 3901 GAACAAATTCAAACAGAGGCAATCGGGTGTGTCGCAACAGCACCAGCAGGAGGCT 3960  
Db  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTCGGGGAGCATATG 4020  
Db  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTCGGGGAGCATATG 4020  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCTCTGCTTGGCAACCCC 4080  
Db  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCTCTGCTTGGCAACCCC 4080  
QY 4081 GCGATAGCATCATGATGGCATTCACAGCCTTATACACAGCCTGATACACCCCAACAT 4140  
Db  
QY 4081 GCGATAGCATCATGATGGCATTCACAGCCTTATACACAGCCTGATACACCCCAACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTCGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
Db  
QY 4141 ACCCTCTGTTTAAACATCTCGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTCTTTTCTGAGCGCGCATCGCTGGAGCGCTGTGTGGCAGCATAGGCTTTGGG 4260  
Db  
QY 4261 AAGTGTCTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGGCTGCTGGCC 4320  
Db  
QY 4261 AAGTGTCTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGGCTGCTGGCC 4320  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT 4380  
Db  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGCGCCCTAGTCTGCGGGTGTGTGGCAGCATGCTGCGTGGCAC 4440  
Db  
QY 4441 GTGGGCCAGGAGGAGGGGCTGTGCACTGATGAA CCGGCTGATAGCGTTTCCGTTCCGGG 4500  
Db  
QY 4441 GTGGGCCAGGAGGAGGGGCTGTGCACTGATGAA CCGGCTGATAGCGTTTCCGTTCCGGG 4500  
QY 4501 GGTAAACACGCTCTCCCCCAGCATATGTGCTGAGAGCGGCTGCGACGCTGCTCACT 4560



```
Qy 6721 CCCGAAGCCAGACAGGCGATAGGCTCGCTCAGAGCGGGCTTTTACATCGGGGGCCCCCTG 6780
Db 6721 CCCGAAGCCAGACAGGCGATAGGCTCGCTCAGAGCGGGCTTTTACATCGGGGGCCCCCTG 6780
Qy 6781 ACTAATTTCTAAAGGCGCAAACTGGGGCTATCGCGGTGTCGCCGGCGAGCGGTGTACTGACG 6840
Db 6781 ACTAATTTCTAAAGGCGCAAACTGGGGCTATCGCGGTGTCGCCGGCGAGCGGTGTACTGACG 6840
Qy 6841 ACCAGCTGCGGTAAATACCCCTCACATGTGTTACTTTGAAGCGCGTGTGCGGCTGTGCGAGCTGCG 6900
Db 6841 ACCAGCTGCGGTAAATACCCCTCACATGTGTTACTTTGAAGCGCGTGTGCGGCTGTGCGAGCTGCG 6900
Qy 6901 AAGCTCCAGGACTCGAGATCGCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAGC 6960
Db 6901 AAGCTCCAGGACTCGAGATCGCTCGTATGCGGAGACGACCTTGTGTTATCTGTGAAGC 6960
Qy 6961 GCGGGGACCCAAAGAGGACGAGCGAGCTTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020
Db 6961 GCGGGGACCCAAAGAGGACGAGCGAGCTTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020
Qy 7021 TCTGCCCCCTGGGACCCGCCCAAAACAGAAATGAGTGGAGTTGATTAACATCATGC 7080
Db 7021 TCTGCCCCCTGGGACCCGCCCAAAACAGAAATGAGTGGAGTTGATTAACATCATGC 7080
Qy 7081 TCCTCCATGTGTAGTCGCGCAGATGATCTGCGCAAAAGGTTGACTATCTCACCGT 7140
Db 7081 TCCTCCATGTGTAGTCGCGCAGATGATCTGCGCAAAAGGTTGACTATCTCACCGT 7140
Qy 7141 GACCCCAACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
Db 7141 GACCCCAACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
Qy 7201 TCCTGGCTAGGCAACATCATGATGCGCCACCTTGTGGGCAAGGATGATCCTGATG 7260
Db 7201 TCCTGGCTAGGCAACATCATGATGCGCCACCTTGTGGGCAAGGATGATCCTGATG 7260
Qy 7261 ACTCATTTCTTCTCCATCCTTCTAGCTCAGAAACAATTTGAAAAGCCCTAGATTGTCAG 7320
Db 7261 ACTCATTTCTTCTCCATCCTTCTAGCTCAGAAACAATTTGAAAAGCCCTAGATTGTCAG 7320
Qy 7321 ATCTACGGGGCTGTTACTCCTATTGAGCACTTACCTTACCTCAGATCAATCAACGACTC 7380
Db 7321 ATCTACGGGGCTGTTACTCCTATTGAGCACTTACCTTACCTCAGATCAATCAACGACTC 7380
Qy 7381 CATGGCTTAGCGATTTTACTCCATGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
Db 7381 CATGGCTTAGCGATTTTACTCCATGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
Qy 7441 TCATGCTCAGGAACCTGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500
Db 7441 TCATGCTCAGGAACCTGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500
Qy 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC 7560
Db 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC 7560
Qy 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGTCCAGTTGGAT 7620
Db 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGTCCAGTTGGAT 7620
Qy 7621 TTATCCAGCTGGTTCGTTGCTGTGTATGAGCGGGGAGACATATATCACAGCTGTCTCGT 7680
Db 7621 TTATCCAGCTGGTTCGTTGCTGTGTATGAGCGGGGAGACATATATCACAGCTGTCTCGT 7680
Qy 7681 GCCCGACCCCGCTGTTTCATGTGTGCTACTCTTACTTTCTGTAGGGGTAGGATCTAT 7740
Db 7681 GCCCGACCCCGCTGTTTCATGTGTGCTACTCTTACTTTCTGTAGGGGTAGGATCTAT 7740
Qy 7741 CTACTCCCAACCGATGAACGGGAGCTTAAACACTTCCAGGCCAATAGGCCATCTCTGTTT 7800
Db 7741 CTACTCCCAACCGATGAACGGGAGCTTAAACACTTCCAGGCCAATAGGCCATCTCTGTTT 7800
```

```
Qy 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Qy 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
Db 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
Qy 7921 TAGCTGTGAAGGTCCTGTGAGCCGCTTGACTGCGAGAGTCTGATATCGGCTTCTTGC 7980
Db 7921 TAGCTGTGAAGGTCCTGTGAGCCGCTTGACTGCGAGAGTCTGATATCGGCTTCTTGC 7980
Qy 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 2
US-10-005-469-1
; Sequence 1, Application US/10005469
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPI
; FILE REFERENCE: 0342/1H395US1
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV replicon 1377/NS3-3'UTR
US-10-005-469-1

Query Match 100.0%; Score 7987.4; DB 13; Length 7992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCAGCCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGAACTACTG 60
Db 1 GCCAGCCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGAACTACTG 60
Qy 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGAGCTCCAGAC 120
Db 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGAGCTCCAGAC 120
Qy 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGCG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGCG 180
Qy 181 GACGACCGGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 240
Db 181 GACGACCGGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 240
Qy 241 GCGAGACTGTAGCCGAGTGTGTGGTTCGGAAGGCCCTTGTGTTACTGCTCTGATAGG 300
Db 241 GCGAGACTGTAGCCGAGTGTGTGGTTCGGAAGGCCCTTGTGTTACTGCTCTGATAGG 300
Qy 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCCTTAAAC 360
Db 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCCTTAAAC 360
Qy 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTC 420
Db 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTC 420
Qy 421 CGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACACAGCAATCGGCTGCT 480
```



Db	421	CGSCCGCTTGGGTGGAGAGCTATTTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCT	480	Db	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAAGCCAGCTGTATAGATACACCT	1560
Qy	481	CTGATCGCGCGTGTTCGGCTGTGAGCGAGGGGCGCCGGTTCCTTTTGTCAAGACCG	540	Qy	1561	GCAAGGGCGGCAAAACCCAGTGCACACCTTGTGAGTTGATAGTTGGAAGAGTCAAA	1620
Db	481	CTGATCGCGCGTGTTCGGCTGTGAGCGAGGGGCGCCGGTTCCTTTTGTCAAGACCG	540	Db	1561	GCAAGGGCGGCAAAACCCAGTGCACACCTTGTGAGTTGATAGTTGGAAGAGTCAAA	1620
Qy	541	ACCTGTCCGGTGCCTGAAATGAATCTCAGACAGAGCAGCGCGGCTATCGTGGCTGGCCA	600	Qy	1621	TGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGATGCCAGAAAGGTACCCCAATGT	1680
Db	541	ACCTGTCCGGTGCCTGAAATGAATCTCAGACAGAGCAGCGCGGCTATCGTGGCTGGCCA	600	Db	1621	TGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGATGCCAGAAAGGTACCCCAATGT	1680
Qy	601	CGACGGGCGCTTCCTTGGCAGCTGTGCTCGACCTTGTCACTGAAGCGGGAAGGCACTGGC	660	Qy	1681	ATGGGATCTGATCTGGGGGCTCGGTGCACATGCTTTTACATGCTTTAGTCGAGGTTAAAA	1740
Db	601	CGACGGGCGCTTCCTTGGCAGCTGTGCTCGACCTTGTCACTGAAGCGGGAAGGCACTGGC	660	Db	1681	ATGGGATCTGATCTGGGGGCTCGGTGCACATGCTTTTACATGCTTTAGTCGAGGTTAAAA	1740
Qy	661	TGCTATTGGCGAAGTGCAGGGGCGAGATCTCTGTCACTGATCTGCTGCGGAGA	720	Qy	1741	AAGCTTAGCCCCCGAACCACCGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Db	661	TGCTATTGGCGAAGTGCAGGGGCGAGATCTCTGTCACTGATCTGCTGCGGAGA	720	Db	1741	AAGCTTAGCCCCCGAACCACCGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Qy	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGATCCGGCTACCTGCC	780	Qy	1801	ATGGCGCTATTACGGCCTACTCCCAACAGACCGGAGGCTACTTGGCTGCATCATCACT	1860
Db	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGATCCGGCTACCTGCC	780	Db	1801	ATGGCGCTATTACGGCCTACTCCCAACAGACCGGAGGCTACTTGGCTGCATCATCACT	1860
Qy	781	CAFTTCAGACCAACGAAACATTCGCATCGAGCGAGCAGCTACTCGGATGGAAAGCCGGTC	840	Qy	1861	AGCTCACAGGCGGGAACAGGAAACAGGTTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
Db	781	CAFTTCAGACCAACGAAACATTCGCATCGAGCGAGCAGCTACTCGGATGGAAAGCCGGTC	840	Db	1861	AGCTCACAGGCGGGAACAGGAAACAGGTTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
Qy	841	TTGTGATCAGATGATCTGGAGAAAGATCGAGGAGCATCAGGGGCTCGCGCCAGCGAACTGTTGC	900	Qy	1921	ACAAATCTTCTGCGGACCTGCGTCAATGGCGTGTGGACTGTGTGACTGTCTATCATGTGCC	1980
Db	841	TTGTGATCAGATGATCTGGAGAAAGATCGAGGAGCATCAGGGGCTCGCGCCAGCGAACTGTTGC	900	Db	1921	ACAAATCTTCTGCGGACCTGCGTCAATGGCGTGTGGACTGTGTGACTGTCTATCATGTGCC	1980
Qy	901	CAAGGCTCAAGCGCGCATCGCGAGGAGCATCGTGTGAGTCACTGAGTTCGAGTGGCGGC	960	Qy	1981	GGCTCAAGAACCTTCCCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Db	901	CAAGGCTCAAGCGCGCATCGCGAGGAGCATCGTGTGAGTCACTGAGTTCGAGTGGCGGC	960	Db	1981	GGCTCAAGAACCTTCCCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Qy	961	GCTTGCAGAAATCATGTTGGAAATGGCGCTTTCTGGATTCACTGAGTTCGAGTGGCGGC	1020	Qy	2041	CAGGACCTCGTGGTGGCAAGCGCCCCCGGGCGCTTCTCTTGAACCATGCACTGTC	2100
Db	961	GCTTGCAGAAATCATGTTGGAAATGGCGCTTTCTGGATTCACTGAGTTCGAGTGGCGGC	1020	Db	2041	CAGGACCTCGTGGTGGCAAGCGCCCCCGGGCGCTTCTCTTGAACCATGCACTGTC	2100
Qy	1021	TGGGTGTGGGACCCCTATCAGGACATAGCGTTGGCTACCGCTGATATGCTGAAGAC	1080	Qy	2101	GGCAGCTCGGACTTTACTTGGTCACGAGGCAATGCGATGTCATTCGGGTGCGCGCGG	2160
Db	1021	TGGGTGTGGGACCCCTATCAGGACATAGCGTTGGCTACCGCTGATATGCTGAAGAC	1080	Db	2101	GGCAGCTCGGACTTTACTTGGTCACGAGGCAATGCGATGTCATTCGGGTGCGCGCGG	2160
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCCG	1140	Qy	2161	GGCGACAGCAGGGGAGCGCTACTCTCCCGCAGCGCGCTCTCTTCTTGAAGGCTCTTCG	2220
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCCG	1140	Db	2161	GGCGACAGCAGGGGAGCGCTACTCTCCCGCAGCGCGCTCTCTTCTTGAAGGCTCTTCG	2220
Qy	1141	AGCGCATCGCTTCTATCGCTTCTTTCAGAGTTCCTTCTGAGTTTAAACAGACCAACAG	1200	Qy	2221	GGCGTCCACTGCTGCTGCGGCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGTGTC	2280
Db	1141	AGCGCATCGCTTCTATCGCTTCTTTCAGAGTTCCTTCTGAGTTTAAACAGACCAACAG	1200	Db	2221	GGCGTCCACTGCTGCTGCGGCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGTGTC	2280
Qy	1201	GTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCTTACGTTACTGCG	1260	Qy	2281	ACCGAGGGGTTGCGAAGGCGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG	2340
Db	1201	GTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCTTACGTTACTGCG	1260	Db	2281	ACCGAGGGGTTGCGAAGGCGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG	2340
Qy	1261	CGAAGCGCTTGAATAAGCGGCTGCTGCTTGTCTATATGTTATTTTCCACCATATG	1320	Qy	2341	CGGTCCCGGCTTCTTACCGCAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTTG	2400
Db	1261	CGAAGCGCTTGAATAAGCGGCTGCTGCTTGTCTATATGTTATTTTCCACCATATG	1320	Db	2341	CGGTCCCGGCTTCTTACCGCAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTTG	2400
Qy	1321	CGCTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTTCAGCAGCATTCCT	1380	Qy	2401	GCCCATCTACGCCCTACTTGTAGCGGCAAGAGCACTAAGGTGCGCGTGGTATGCA	2460
Db	1321	CGCTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTTCAGCAGCATTCCT	1380	Db	2401	GCCCATCTACGCCCTACTTGTAGCGGCAAGAGCACTAAGGTGCGCGTGGTATGCA	2460
Qy	1381	AGGGGCTTCTTCCCTCTCGCCAAAGGAATGCAAGGCTGTGTAATGTGCTGAAGGAGCA	1440	Qy	2461	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGCTCGCGCCACCTTAGGTTTCGGG	2520
Db	1381	AGGGGCTTCTTCCCTCTCGCCAAAGGAATGCAAGGCTGTGTAATGTGCTGAAGGAGCA	1440	Db	2461	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGCTCGCGCCACCTTAGGTTTCGGG	2520
Qy	1441	GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTTCAGCGCAGCGG	1500	Qy	2521	GGGTATATCTTAAGGCACATGGTATCGACCCCTAACATCAGAACCGGGGTAAAGCACCATC	2580
Db	1441	GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTTCAGCGCAGCGG	1500	Db	2521	GGGTATATCTTAAGGCACATGGTATCGACCCCTAACATCAGAACCGGGGTAAAGCACCATC	2580
Qy	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAAGCCAGCTGTATAGATACACCT	1560	Qy	2581	ACCAAGGTTGCCCGGCTCGTACCTACCTATGGCAAGTTCCTTCCGACGGTGGTTC	2640
Db	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAAGCCAGCTGTATAGATACACCT	1560	Db	2581	ACCAAGGTTGCCCGGCTCGTACCTACCTATGGCAAGTTCCTTCCGACGGTGGTTC	2640

QY 2641 TCTGGGGCGCCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGTGGGCATCGGCACAGCTCTTGACCAAGCGGAGCGGCTGAGCGCGACTCGTCGTG 2760  
Db 2701 ATCTGTGGGCATCGGCACAGCTCTTGACCAAGCGGAGCGGCTGAGCGCGACTCGTCGTG 2760  
QY 2761 CTGCGCCACCGCTACGCTCCGGGATCGGTCAACGCTGCGACATCAAAATCGAGGAGGTG 2820  
Db 2761 CTGCGCCACCGCTACGCTCCGGGATCGGTCAACGCTGCGACATCAAAATCGAGGAGGTG 2820  
QY 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGCGCAAGCAATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGCGCAAGCAATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAAATGTGATGAGCTCGCCGCG 2940  
Db 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAAATGTGATGAGCTCGCCGCG 2940  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGTCTGTAGCATATTAACGGGGCTTGATGTATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGTCTGTAGCATATTAACGGGGCTTGATGTATCCGTC 3000  
QY 3001 ATACCAACTAGCGAGAGCGTCAATTTCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGAGAGCGTCAATTTCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
QY 3061 GCGGATTTGCACTCAGTGATCGAATGCAATACATGTGTCAACAGACAGTCGACTTCAGC 3120  
Db 3061 GCGGATTTGCACTCAGTGATCGAATGCAATACATGTGTCAACAGACAGTCGACTTCAGC 3120  
QY 3121 CTGGAACCGCACTTCAATTTGAGACGACGACCGTGTGCAACAGCGGGTGTACGCTCG 3180  
Db 3121 CTGGAACCGCACTTCAATTTGAGACGACGACCGTGTGCAACAGCGGGTGTACGCTCG 3180  
QY 3181 CAGCGGCGAGGACGAGTGTGTAGGGGAGGATGGGCATTTACAGGTTGTGACTCCAGGA 3240  
Db 3181 CAGCGGCGAGGACGAGTGTGTAGGGGAGGATGGGCATTTACAGGTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGCTGT 3300  
QY 3301 GCTTGTACGAGTCAACGCGCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTACGAGTCAACGCGCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGGCTC 3420  
Db 3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGGCTC 3420  
QY 3421 ACCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGACAACTTCCCTTAC 3480  
Db 3421 ACCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGACAACTTCCCTTAC 3480  
QY 3481 CTGTTAGCATACAGGCTACGTTGTGCGCCAGGCTCAGGCTCCACTCCATCTGTGGAC 3540  
Db 3481 CTGTTAGCATACAGGCTACGTTGTGCGCCAGGCTCAGGCTCCACTCCATCTGTGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTAGCTGACGCGCCCAACGCCCTCTGT 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTAGCTGACGCGCGCAACGCCCTCTGT 3600  
QY 3601 TATAGGCTGGAGCGCTTCAAAACGAGTTTACTACCAACACACACCCCAATAACCAATACATC 3660  
Db 3601 TATAGGCTGGAGCGCTTCAAAACGAGTTTACTACCAACACACACCCCAATAACCAATACATC 3660  
QY 3661 ATGSCATGCAATGCTGGCTGACCTGGAGTGTCTGACGAGCACTGGGTGCTGGTAGGGGA 3720  
Db 3661 ATGSCATGCAATGCTGGCTGACCTGGAGTGTCTGACGAGCACTGGGTGCTGGTAGGGGA 3720

QY 3721 GTCTAGCAGCTCTGGCGCGGTATTGCTGACAAACAGGACGCTGTTCAATGTGGCGAGG 3780  
Db 3721 GTCTAGCAGCTCTGGCGCGGTATTGCTGACAAACAGGACGCTGTTCAATGTGGCGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAAGCGGCGCATCATTTCCCGACAGGGAAGTCCTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAAGCGGCGCATCATTTCCCGACAGGGAAGTCCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGGCTCTACACTCCCTTTACATCGAACAGGGAATCAGCTCGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGCTCTACACTCCCTTTACATCGAACAGGGAATCAGCTCGCC 3900  
QY 3901 GAACAAATTCACACAGAGGCAATCGGGTGTCTGCAACAGCCACCAACAGGAGCGGCT 3960  
Db 3901 GAACAAATTCACACAGAGGCAATCGGGTGTCTGCAACAGCCACCAACAGGAGCGGCT 3960  
QY 3961 GCTGCTCCCGTGTGGAAATCCAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAACAATG 4020  
Db 3961 GCTGCTCCCGTGTGGAAATCCAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAACAATG 4020  
QY 4021 TGGAAATTTCACTAGCGGGATACAAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTTCACTAGCGGGATACAAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCATTCACAGCCCTTATACAGCCCTTATACAGCCCGCTCACACCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCATTCACAGCCCTTATACAGCCCTTATACAGCCCGCTCACACCACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTCTTGTGTAGCGCGCGCATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTCTTGTGTAGCGCGCGCATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGG 4260  
QY 4261 AAGTGCTCTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGGCTCGTGCGC 4320  
Db 4261 AAGTGCTCTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGGCTCGTGCGC 4320  
QY 4321 TTTAAGGTCATAGCGCGGAGATGCCCTCCACCGAGCACTGTGTTAACTTACTCTCCCTGCT 4380  
Db 4321 TTTAAGGTCATAGCGCGGAGATGCCCTCCACCGAGCACTGTGTTAACTTACTCTCCCTGCT 4380  
QY 4381 ATCTCTTCCCTGCGCCCTAGTGTCTCGGGTGTGTGTGCGAGCGATATGCTGCTGGGAC 4440  
Db 4381 ATCTCTTCCCTGCGCCCTAGTGTCTCGGGTGTGTGTGCGAGCGATATGCTGCTGGGAC 4440  
QY 4441 GTGGGCGCGAGGAGGGGCTGTGCACTGGATGAACCGGCTGTAGGTTGCTTCCGG 4500  
Db 4441 GTGGGCGCGAGGAGGGGCTGTGCACTGGATGAACCGGCTGTAGGTTGCTTCCGG 4500  
QY 4501 GGTAAACCACTCTCTCCCGCAGCGCATATGTGCTGAGAGCGACGCTGAGCAGCTGTCACT 4560  
Db 4501 GGTAAACCACTCTCTCCCGCAGCGCATATGTGCTGAGAGCGACGCTGAGCAGCTGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTTAACAATCACTCAGCTGTCTGAAGAGCTTCCAGTGTGATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTTAACAATCACTCAGCTGTCTGAAGAGCTTCCAGTGTGATCAAC 4620  
QY 4621 GAGGACTGTCCAGCGCATCTCGGCTCGTGGCTAGAGATGTTTGGGATTTGATATGC 4680  
Db 4621 GAGGACTGTCCAGCGCATCTCGGCTCGTGGCTAGAGATGTTTGGGATTTGATATGC 4680  
QY 4681 ACGGTGTGACTGATTTTCAAGACTGCTGCTCAGTCCAGGCTCTCTGCGCGCATTTGCCGGGA 4740  
Db 4681 ACGGTGTGACTGATTTTCAAGACTGCTGCTCAGTCCAGGCTCTCTGCGCGCATTTGCCGGGA 4740  
QY 4741 GTCCCTTCTTCTCATGTCAACGTTGGGTACAAAGGAGTCTGGGGGGCGAGCGCATCATG 4800  
Db 4741 GTCCCTTCTTCTCATGTCAACGTTGGGTACAAAGGAGTCTGGGGGGCGAGCGCATCATG 4800  
QY 4801 CAACACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAGG 4860



QY 7021 TCTGCCCCCTGGGGCCCGCCAAACAGAAATACGACTTGGAGTTGATACATATGC 7080  
 Db 7021 TCTGCCCCCTGGGGCCCGCCAAACAGAAATACGACTTGGAGTTGATACATATGC 7080  
 QY 7081 TCCTCCAAATGTGTCAGTCGGCCACGATCATCTGCAAAAGGGTGTACTATCTCACCGGT 7140  
 Db 7081 TCCTCCAAATGTGTCAGTCGGCCACGATCATCTGCAAAAGGGTGTACTATCTCACCGGT 7140  
 QY 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 Db 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 QY 7201 TCCTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGGATCATCTGATG 7260  
 Db 7201 TCCTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGGATCATCTGATG 7260  
 QY 7261 ACTCATTTCTTCTCATCTCTAGCTCAGAAACAACTTGAAGAGCCCTAGATTGTGAC 7320  
 Db 7261 ACTCATTTCTTCTCATCTCTAGCTCAGAAACAACTTGAAGAGCCCTAGATTGTGAC 7320  
 QY 7321 ATCTACGGGGCTGTACTCATTTAGGCCACTTGCACCTTACCTACCTCAGATCAATCAACGACTC 7380  
 Db 7321 ATCTACGGGGCTGTACTCATTTAGGCCACTTGCACCTTACCTACCTCAGATCAATCAACGACTC 7380  
 QY 7381 CATGSCCTTAGGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
 Db 7381 CATGSCCTTAGGCAATTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
 QY 7441 TCATGSCCTCAGAAACCTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAT 7500  
 Db 7441 TCATGSCCTCAGAAACCTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAT 7500  
 QY 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
 Db 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
 QY 7561 AACTGGGCAGTAAGCAACAGCTCAACTCACTCCACTCCAGCTGCTGCCAGTTGAT 7620  
 Db 7561 AACTGGGCAGTAAGCAACAGCTCAACTCACTCCACTCCAGCTGCTGCCAGTTGAT 7620  
 QY 7621 TTATCCAGCTGGTTCGTGGTTACAGCGGGGAGACATATATCACAGCTGCTCTCGT 7680  
 Db 7621 TTATCCAGCTGGTTCGTGGTTACAGCGGGGAGACATATATCACAGCTGCTCTCGT 7680  
 QY 7681 GCCCGACCCCGCTGCTCATGCTGCTGCTACTCTCTACTTCTGTAGGGGTAGGCATCTAT 7740  
 Db 7681 GCCCGACCCCGCTGCTCATGCTGCTGCTACTCTCTACTTCTGTAGGGGTAGGCATCTAT 7740  
 QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTCTTT 7800  
 Db 7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTCTTT 7800  
 QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 QY 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Db 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 QY 7921 TAGCTGTGAAGGTCCTGAGCCGCTTACCTGAGAGTGTCTGATCTGAGAGTGTCTGATCTGCTG 7980  
 Db 7921 TAGCTGTGAAGGTCCTGAGCCGCTTACCTGAGAGTGTCTGATCTGATCTGCTGCTGCTG 7980  
 QY 7981 AGATCAAGT 7989  
 Db 7981 AGATCAAGT 7989

RESULT 3

US-10-434-842-1

; Sequence 1, Application US/10434842

; Publication No. US20040005549A1

; GENERAL INFORMATION:

; APPLICANT: Bichko, Vadim

; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION

; FILE REFERENCE: 0342/1H395US3

; CURRENT APPLICATION NUMBER: US/10/434,842

; CURRENT FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: US 10/233,307

; PRIOR FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: US 10/005,469

; PRIOR FILING DATE: 2001-11-07

; PRIOR APPLICATION NUMBER: US 60/245,866

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 7992

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HCVlb-based chimeric replicon

; US-10-434-842-1

Query Match 100.0%; Score 7987.4; DB 16; Length 7992;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTGGGGGCGACACTCCACATAGATCACTCCCTCTGTGAGGAACTACTG 60  
 Db 1 GCCAGCCCCGATTGGGGGCGACACTCCACATAGATCACTCCCTCTGTGAGGAACTACTG 60  
 QY 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTATGATGAGTGTCTGAGAGCTCCAGAC 120  
 Db 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTATGATGAGTGTCTGAGAGCTCCAGAC 120  
 QY 121 CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGTCTGAGTACACCGAAATGCCAG 180  
 Db 121 CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGTCTGAGTACACCGAAATGCCAG 180  
 QY 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCTCC 240  
 Db 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCTCC 240  
 QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGCGAAGCCCTTGTGTTCTGCTGATAGG 300  
 Db 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGCGAAGCCCTTGTGTTCTGCTGATAGG 300  
 QY 301 GTGCTTGGAGTGCCTCCGGAGGTCTGTAGACCTGCAACCTGAGCAGCAATCTTAAC 360  
 Db 301 GTGCTTGGAGTGCCTCCGGAGGTCTGTAGACCTGCAACCTGAGCAGCAATCTTAAC 360  
 QY 361 CTCAGAGAAACCAAAAGGGCGGCCATGATTGAACAAGATGATTCACGACAGTTCCTC 420  
 Db 361 CTCAGAGAAACCAAAAGGGCGGCCATGATTGAACAAGATGATTCACGACAGTTCCTC 420  
 QY 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGATCTGGGCAACAGCAATTCGCTGCT 480  
 Db 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGATCTGGGCAACAGCAATTCGCTGCT 480  
 QY 481 CTGATCCGCGCTGTTCCGGCTGTACGAGGGCGCCGGTCTTTTCTCAAGACCG 540  
 Db 481 CTGATCCGCGCTGTTCCGGCTGTACGAGGGCGCCGGTCTTTTCTCAAGACCG 540  
 QY 541 ACCTGTCCGCTGCTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 600  
 Db 541 ACCTGTCCGCTGCTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 600  
 QY 601 CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC 660  
 Db 601 CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC 660  
 QY 661 TGCTATTGGGCGAAGTCCCGGGCAGGATCTCCTGTCTCATCTCAGCTTGTCTCTCCGAGA 720

Db	661	TGCTATTGGCGAAGTGCCGGGCGAGGATCTCCTGTCTATCTCACCTTGCTCTCTCCGCGAGA	720
Qy	721	AGATATCCATCATGGGCTGATGCAATGCGGGCTTGATAGGCTTGATCCGGCTACCTGCC	780
Db	721	AGATATCCATCATGGGCTGATGCAATGCGGGCTTGATAGGCTTGATCCGGCTACCTGCC	780
Qy	781	CATTTCGACCAACCAAGCGAACAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCGGGTC	840
Db	781	CATTTCGACCAACCAAGCGAACAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCGGGTC	840
Qy	841	TTGTGATCAGGATGATGGAAGAGAGATCAGAGGCTCGCGCCAGCGGAATCTGTTCG	900
Db	841	TTGTGATCAGGATGATGGAAGAGAGATCAGAGGCTCGCGCCAGCGGAATCTGTTCG	900
Qy	901	CAAGGCTCAAGCGCGCATGCCGAGCGGAGAGATCTCGTCTGACCCATGGGCGATGCCCT	960
Db	901	CAAGGCTCAAGCGCGCATGCCGAGCGGAGAGATCTCGTCTGACCCATGGGCGATGCCCT	960
Qy	961	GCTTGCAGGAATATCATGGTGGAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGC	1020
Db	961	GCTTGCAGGAATATCATGGTGGAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGC	1020
Qy	1021	TGGGTGTGGCGACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGGAAGAGC	1080
Db	1021	TGGGTGTGGCGACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGGAAGAGC	1080
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCCG	1140
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCCG	1140
Qy	1141	AGCGCATCGGCTTCTATCGGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAACG	1200
Db	1141	AGCGCATCGGCTTCTATCGGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAACG	1200
Qy	1201	GTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACGTTACTGGC	1260
Db	1201	GTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACGTTACTGGC	1260
Qy	1261	CGAAGCGGCTTGAATAAGCGCGGTGCGTTGTCTATATGTTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGGCTTGAATAAGCGCGGTGCGTTGTCTATATGTTATTTTCCACCATATTG	1320
Qy	1321	CCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGCTTCTTTGACGAGCATTCCT	1380
Db	1321	CCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGCTTCTTTGACGAGCATTCCT	1380
Qy	1381	AGGGGTCTTCCCTCTCGCAAGGAATGCAAGGCTGTGTTGAATGTCTGTAAGGAAGCA	1440
Db	1381	AGGGGTCTTCCCTCTCGCAAGGAATGCAAGGCTGTGTTGAATGTCTGTAAGGAAGCA	1440
Qy	1441	GTTTCCCTCTGGAAGCTTCTTGAAGCAACAACCTCTGTAGCGACCCCTTTGCGAGCGCG	1500
Db	1441	GTTTCCCTCTGGAAGCTTCTTGAAGCAACAACCTCTGTAGCGACCCCTTTGCGAGCGCG	1500
Qy	1501	AACCCGCCACTCTGGGACAGGTGCTCTGGGCGCAAAAGCCAGCTGTATAAGATACACCT	1560
Db	1501	AACCCGCCACTCTGGGACAGGTGCTCTGGGCGCAAAAGCCAGCTGTATAAGATACACCT	1560
Qy	1561	GCAAGCGGCAACAACCCAGTGCAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGCGGCAACAACCCAGTGCAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
Db	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
Qy	1681	ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAAA	1740
Db	1681	ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAAA	1740
Qy	1741	AAGCTCTAGGCCCCCGAACAACGAGGAGCGTGGTTTCTTTTGTAAAAACACGATATACC	1800
Db	1741	AAGCTCTAGGCCCCCGAACAACGAGGAGCGTGGTTTCTTTTGTAAAAACACGATATACC	1800
Qy	1801	ATGGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCACT	1860
Db	1801	ATGGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCACT	1860
Qy	1861	AGCCTCACAGGCGCGGACAGGAACCAAGTGCAGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCACAGGCGCGGACAGGAACCAAGTGCAGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Qy	1921	ACAAATCTTTCTGGGACCTGCGTCAATGGGCTGTGTGTGGACTGTCTATCATGTGTGCC	1980
Db	1921	ACAAATCTTTCTGGGACCTGCGTCAATGGGCTGTGTGTGGACTGTCTATCATGTGTGCC	1980
Qy	1981	GGCTCAAGACCCCTTCGCGGCCCAAGGCGCCCAATCACCAATGTACCAATGTGGAC	2040
Db	1981	GGCTCAAGACCCCTTCGCGGCCCAAGGCGCCCAATCACCAATGTACCAATGTGGAC	2040
Qy	2041	CAGGACCTCGTGGTGGCAAGCGCCCCCGGGCGCGTTCCTTTGACCAATGCACTGTC	2100
Db	2041	CAGGACCTCGTGGTGGCAAGCGCCCCCGGGCGCGTTCCTTTGACCAATGCACTGTC	2100
Qy	2101	GGCAGCTCGGACCTTTTACTTGTACGAGGCAATGCGATGTCTATTCGGTTCGCGCGCG	2160
Db	2101	GGCAGCTCGGACCTTTTACTTGTACGAGGCAATGCGATGTCTATTCGGTTCGCGCGCG	2160
Qy	2161	GGCGACAGAGGGGAGCCTACTCTCCCGAGGCGCGTCTCTTACTTTGAAGGCTCTTTCG	2220
Db	2161	GGCGACAGAGGGGAGCCTACTCTCCCGAGGCGCGTCTCTTACTTTGAAGGCTCTTTCG	2220
Qy	2221	GGCGGTCCTACTGTCTGCCCTCGGGGCGAGCTGTGGGCTCTTTGGGCTGCGGTGTGC	2280
Db	2221	GGCGGTCCTACTGTCTGCCCTCGGGGCGAGCTGTGGGCTCTTTGGGCTGCGGTGTGC	2280
Qy	2281	ACCGAGGGGTTGCGAAGCGGGTGAGCTTTGTACCGTGTGAGTCTATGGAACCACTATG	2340
Db	2281	ACCGAGGGGTTGCGAAGCGGGTGAGCTTTGTACCGTGTGAGTCTATGGAACCACTATG	2340
Qy	2341	CGGTCCCGGCTTCTACGAGCAACTCTGCTCCCTCGGCGCGTACCGCAGACATTCACAGTG	2400
Db	2341	CGGTCCCGGCTTCTACGAGCAACTCTGCTCCCTCGGCGCGTACCGCAGACATTCACAGTG	2400
Qy	2401	GCCCATCTACGCCCTTACTGGTAGCGGCAAGACACTAAGGTGCGGCTGCGTATGCA	2460
Db	2401	GCCCATCTACGCCCTTACTGGTAGCGGCAAGACACTAAGGTGCGGCTGCGTATGCA	2460
Qy	2461	GCCCAAGGTTAAGGTGCTTGTCTGAAACCGTCTGCGCGCCACCTTAGGTTTCGGG	2520
Db	2461	GCCCAAGGTTAAGGTGCTTGTCTGAAACCGTCTGCGCGCCACCTTAGGTTTCGGG	2520
Qy	2521	GGGTATATGTAAAGCAGATGGTATCGACCTTAAACATCAGAACCGGGTAAAGACCATC	2580
Db	2521	GGGTATATGTAAAGCAGATGGTATCGACCTTAAACATCAGAACCGGGTAAAGACCATC	2580
Qy	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGCAAGTTTCTTCCGACCGTGTGTC	2640
Db	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGCAAGTTTCTTCCGACCGTGTGTC	2640
Qy	2641	TCTTGGGGCGCTTACACATCATATATGTATGAGTGAGTGCACCTCACTGACTGCACT	2700
Db	2641	TCTTGGGGCGCTTACACATCATATATGTATGAGTGAGTGCACCTCACTGACTGCACT	2700
Qy	2701	ATCCTGGGATCGGACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGGCTCGTCTGTC	2760
Db	2701	ATCCTGGGATCGGACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGGCTCGTCTGTC	2760
Qy	2761	CTCGCAACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
Db	2761	CTCGCAACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
Qy	2821	GCTCTCTCAGACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Db	2821	GCTCTCTCAGACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880

2881 AAGGGGGAGGACCTCATTTCTGCCATTCCAAAGAGAAATGTGATGAGCTCCGCGG 2940  
2881 AAGGGGGAGGACCTCATTTCTGCCATTCCAAAGAGAAATGTGATGAGCTCCGCGG 2940  
2941 AAGCTGTCCGGCTCGGACTCAATCTGTAGCATATTAACGGGGCTTTGATGTATCCGTC 3000  
2941 AAGCTGTCCGGCTCGGACTCAATCTGTAGCATATTAACGGGGCTTTGATGTATCCGTC 3000  
3001 ATACCAATAGCGGAGAGCTCATTTCTGTAGCAACGGAGCTCTAATGACGGCTTTTACC 3060  
3001 ATACCAATAGCGGAGAGCTCATTTCTGTAGCAACGGAGCTCTAATGACGGCTTTTACC 3060  
3061 GCGATTTTCGACTCAGTGTGATCGACTGCAATCATGTGTCAACAGACAGTGCACCTTACG 3120  
3061 GCGATTTTCGACTCAGTGTGATCGACTGCAATCATGTGTCAACAGACAGTGCACCTTACG 3120  
3121 CTGGACCCGACTTCAACATTGAGACGACCGGTGCCAAGACGCGGTGTACGCTCG 3180  
3121 CTGGACCCGACTTCAACATTGAGACGACCGGTGCCAAGACGCGGTGTACGCTCG 3180  
3181 CAGCGCGGAGGACGACTGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGGA 3240  
3181 CAGCGCGGAGGACGACTGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGGA 3240  
3241 GAAAGGCGCTCGGGCATGTTCGATTCTCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
3241 GAAAGGCGCTCGGGCATGTTCGATTCTCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
3301 GCTTGGTACGAGCTCACGCGCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
3301 GCTTGGTACGAGCTCACGCGCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTCTCGGAGAGGCTTTTACAGGCTC 3420  
3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTCTCGGAGAGGCTTTTACAGGCTC 3420  
3421 ACCACATPAGAGCGCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
3421 ACCACATPAGAGCGCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
3481 CTGTTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTTCCATCTCGTGGGAC 3540  
3481 CTGTTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTTCCATCTCGTGGGAC 3540  
3541 CAATGTGGAGTGTCTCATAGGCTAAGCTTACGCTGACGGGCAACGCGCTCGTGTG 3600  
3541 CAATGTGGAGTGTCTCATAGGCTAAGCTTACGCTGACGGGCAACGCGCTCGTGTG 3600  
3601 TATAGGCTGGAGCGCTTCAAAAAGAGTTACTACACACACCCCATAAACCAATACATC 3660  
3601 TATAGGCTGGAGCGCTTCAAAAAGAGTTACTACACACACCCCATAAACCAATACATC 3660  
3661 ATGGCATGCAATGTCCGGTGAACCTGGAGTGTCTACAGACACCTGGGTCTGTAGGCGGA 3720  
3661 ATGGCATGCAATGTCCGGTGAACCTGGAGTGTCTACAGACACCTGGGTCTGTAGGCGGA 3720  
3721 GTCCTAGCACTCTGGCGCGTATGCTGTGCAACAGCAGCAGTGTGATTTGTGGGAGG 3780  
3721 GTCCTAGCACTCTGGCGCGTATGCTGTGCAACAGCAGCAGTGTGATTTGTGGGAGG 3780  
3781 ATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAAGTCTTTTACCGGGAGTTC 3840  
3781 ATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGGAAGTCTTTTACCGGGAGTTC 3840  
3841 GATGAGATGGAAGTGTCCGCTTCAACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
3841 GATGAGATGGAAGTGTCCGCTTCAACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
3901 GAACAATTCACACAGAGGCAATCGGTTGTGTGCAACAGCCACACAGCAAGCGGAGGCT 3960  
3901 GAACAATTCACACAGAGGCAATCGGTTGTGTGCAACAGCCACACAGCAAGCGGAGGCT 3960  
3961 GCTGCTCCCGGTGGGAATCCAAGTGGCGGACCTTCGGAAGCTTCTGGGCGAAGCATATG 4020

3961 GCTGCTCCCGGTGGGAATCCAAGTGGCGGACCTTCGAAGCCCTTCTGGGCGAAGCATATG 4020  
4021 TGGAAATTTTCATCAGCGGATACAATATTTAGCAGGCTTTGTCCACTCTGCTGGCAACCC 4080  
4021 TGGAAATTTTCATCAGCGGATACAATATTTAGCAGGCTTTGTCCACTCTGCTGGCAACCC 4080  
4081 GCGATAGCATCACTGATGGCAATTCACAGCCTTATCACAGCCCGCTCACACCAACAT 4140  
4081 GCGATAGCATCACTGATGGCAATTCACAGCCTTATCACAGCCCGCTCACACCAACAT 4140  
4141 ACCCTCTGTTTAAATCCTCGGGGATGGGTGGCCGCCCACTTGTCTCTCCAGCGCT 4200  
4141 ACCCTCTGTTTAAATCCTCGGGGATGGGTGGCCGCCCACTTGTCTCTCCAGCGCT 4200  
4201 GCTTCTGCTTTCTGAGGCGCGCATCTGCTGAGCGGCTGTTGGCAGCATAGGCTTTGG 4260  
4201 GCTTCTGCTTTCTGAGGCGCGCATCTGCTGAGCGGCTGTTGGCAGCATAGGCTTTGG 4260  
4261 AAGTGTCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCTCTGTGCCC 4320  
4261 AAGTGTCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCTCTGTGCCC 4320  
4321 TTTAAGGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCTCTGCT 4380  
4321 TTTAAGGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCTCTGCT 4380  
4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGGTCTGTGCGCAGCGATATCTGCTCGGCAC 4440  
4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGGTCTGTGCGCAGCGATATCTGCTCGGCAC 4440  
4441 GTGGGCCAGGGAGGGGCTGTGCAAGTGTGATGAACCGGCTGTAGCGTTTCGTTTCGGG 4500  
4441 GTGGGCCAGGGAGGGGCTGTGCAAGTGTGATGAACCGGCTGTAGCGTTTCGTTTCGGG 4500  
4501 GGTAAACACAGCTCTCCCGACGACTATGTGCTCAGAGCGAGCTGCGAGCGTGTCACT 4560  
4501 GGTAAACACAGCTCTCCCGACGACTATGTGCTCAGAGCGAGCTGCGAGCGTGTCACT 4560  
4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGGCTTTCACAGTGGATCAAC 4620  
4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGGCTTTCACAGTGGATCAAC 4620  
4621 GAGACTGTCTCCAGCGCATGCTCCGGCTCTGGGTCTGAAGATGTTTGGGATTTGATATGC 4680  
4621 GAGACTGTCTCCAGCGCATGCTCCGGCTCTGGGTCTGAAGATGTTTGGGATTTGATATGC 4680  
4681 ACGGTGTTGACTGATTTCAAGACCTGGTCCAGTCCAGTCTCCGCGGATTCGCGGGA 4740  
4681 ACGGTGTTGACTGATTTCAAGACCTGGTCCAGTCTCCGCGGATTCGCGGGA 4740  
4741 GTCCTCTTCTCTCATGTCACAGTGGGTCAAGGAGTCTGCGCGGGCGAGCGCATCATG 4800  
4741 GTCCTCTTCTCTCATGTCACAGTGGGTCAAGGAGTCTGCGCGGGCGAGCGCATCATG 4800  
4801 CAACACACCTGCTCCATGTGGAGCAAGATCAACCGGATGTAAGGATTTCCATGAGG 4860  
4801 CAACACACCTGCTCCATGTGGAGCAAGATCAACCGGATGTAAGGATTTCCATGAGG 4860  
4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAAGCGGTAC 4920  
4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAAGCGGTAC 4920  
4921 ACCACGGGCTTCTGACGCGCTTCCCGGCGCAAAATTTATTTAGGCGGCTGTGGCGGCTG 4980  
4921 ACCACGGGCTTCTGACGCGCTTCCCGGCGCAAAATTTATTTAGGCGGCTGTGGCGGCTG 4980  
4981 GCTGCTGAGGAGTACGTGGAGGTTTACCGGGTGGGATTTTCCACTACGTGACGGGATG 5040  
4981 GCTGCTGAGGAGTACGTGGAGGTTTACCGGGTGGGATTTTCCACTACGTGACGGGATG 5040  
5041 ACCACTCACAACTGTAAGTCCCGTGTCAAGTTCGCGGCCCGCAATTTCTTACAGAGTG 5100









Db 961 GCTTCCGAATATCATGTGTGAAATGGCCGCTTTCTCGAATCATCGACTGTGGCCGC 1020  
QY 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTGTGGCTACCGTGATATGTGTAAGAGC 1080  
Db 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTGTGGCTACCGTGATATGTGTAAGAGC 1080  
QY 1081 TTGGCGGCAATGGGCTGACCGCTTCCCTCGTGTGTTTACGGTATCGCGCTCCCGATTGCG 1140  
Db 1081 TTGGCGGCAATGGGCTGACCGCTTCCCTCGTGTGTTTACGGTATCGCGCTCCCGATTGCG 1140  
QY 1141 AGCGCATCGCCTTCTATCGCCTTCTTACGAGTTCCTTCTGAGTTTAAACAGACCAACAAG 1200  
Db 1141 AGCGCATCGCCTTCTATCGCCTTCTTACGAGTTCCTTCTGAGTTTAAACAGACCAACAAG 1200  
QY 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTAACGTTACTGGC 1260  
Db 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCGCCCTAACGTTACTGGC 1260  
QY 1261 CGAAGCGGCTGGAATAAGCGGCTGTGCTGTGCTATATGTTATTTTCCACCATATTG 1320  
Db 1261 CGAAGCGGCTGGAATAAGCGGCTGTGCTGTGCTATATGTTATTTTCCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGTCTTCTTGACGAGCAATCCT 1380  
Db 1321 CCGTCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGTCTTCTTGACGAGCAATCCT 1380  
QY 1381 AGGGTCTTTCCCTCTCCGCAAGGAATGCAAGGCTGTGATGTGTGTAAGGAAGCA 1440  
Db 1381 AGGGTCTTTCCCTCTCCGCAAGGAATGCAAGGCTGTGATGTGTGTAAGGAAGCA 1440  
QY 1441 GTTCTCTGAGAGCTTTTGAAGACAAACAAGCTGTGAGGACCCCTTTGACGAGCGG 1500  
Db 1441 GTTCTCTGAGAGCTTTTGAAGACAAACAAGCTGTGAGGACCCCTTTGACGAGCGG 1500  
QY 1501 AACCCGCCCTGGGACAGGTGCTCTGCGGCCAAAGCCACGCTGTATAAGATACACCT 1560  
Db 1501 AACCCGCCCTGGGACAGGTGCTCTGCGGCCAAAGCCACGCTGTATAAGATACACCT 1560  
QY 1561 GCAAGGGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
Db 1561 GCAAGGGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGCGCTCTCTCAAGCGTATTCAACAGGGGCTGCAAGGATGCCAGAGGTACCCCATTTG 1680  
Db 1621 TGCGCTCTCTCAAGCGTATTCAACAGGGGCTGCAAGGATGCCAGAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTTTGTAGTCGAGGTTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTTTGTAGTCGAGGTTAAAA 1740  
QY 1741 AACGTCTAGGCCCCCGAACACCGGGAGCTGGTTTCTTTGAAAAACACGATATACC 1800  
Db 1741 AACGTCTAGGCCCCCGAACACCGGGAGCTGGTTTCTTTGAAAAACACGATATACC 1800  
QY 1801 ATGGGCGCTATTACGCGCTACTCCCAACAGCGGAGGCTACTTCGCTGCATCATCACT 1860  
Db 1801 ATGGGCGCTATTACGCGCTACTCCCAACAGCGGAGGCTACTTCGCTGCATCATCACT 1860  
QY 1861 AGCCTCACAGCGCGGACAGGAACACAGGTTCGAGGGGAGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCCTCACAGCGCGGACAGGAACACAGGTTCGAGGGGAGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGGGACCTGCTGCTCAATGGGCTGTGTTGATCTGATCATGTGGC 1980  
Db 1921 ACACATCTTTCTGGGACCTGCTGCTCAATGGGCTGTGTTGATCTGATCATGTGGC 1980  
QY 1981 GGCTCAAGACCTTCCCGGCCCAAGAGGCCCAATCAACCAATGTACACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCTTCCCGGCCCAAGAGGCCCAATCAACCAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGTGACCAATGACCTGCG 2100

Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGTGACCAATGACACTGC 2100  
QY 2101 GCGAGCTCGGACCTTTACTTGTTCAGGAGCATGCGATGTATTCGCTGCGCCGCGG 2160  
Db 2101 GCGAGCTCGGACCTTTACTTGTTCAGGAGCATGCGATGTATTCGCTGCGCCGCGG 2160  
QY 2161 GCGCAGCAGGAGGAGCCTACTCTCCCGCAGCGCCGCTCTCTACTTTGAAGGCTCTTGC 2220  
Db 2161 GCGCAGCAGGAGGAGCCTACTCTCTCCCGCAGCGCCGCTCTCTACTTTGAAGGCTCTTGC 2220  
QY 2221 GCGGCTCCACTGCTCTGCCCCGCGGACAGGTGTGGGCACTTTTCGCGCTGCGGTGTC 2280  
Db 2221 GCGGCTCCACTGCTCTGCCCCGCGGACAGGTGTGGGCACTTTTCGCGCTGCGGTGTC 2280  
QY 2281 ACCGAGGGGTTGCGAAGCGGTGACTTTGTACCGCTGAGTCTATGTGAAGAACACTATG 2340  
Db 2281 ACCGAGGGGTTGCGAAGCGGTGACTTTGTACCGCTGAGTCTATGTGAAGAACACTATG 2340  
QY 2341 CCGTCCCCCGCTTTTCAAGGACAACTCGTCCCCCTCCGCGCGTACCAGAGCACTTCCAGGTG 2400  
Db 2341 CCGTCCCCCGCTTTTCAAGGACAACTCGTCCCCCTCCGCGCGTACCAGAGCACTTCCAGGTG 2400  
QY 2401 GCGCATCTACACGCGCTACTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
Db 2401 GCGCATCTACACGCGCTACTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
QY 2461 GCGCAGGGTATAAGGTGCTTCTGTAACCGCTGCTGCGCGCACCCCTAGGTTTCGGG 2520  
Db 2461 GCGCAGGGTATAAGGTGCTTCTGTAACCGCTGCTGCGCGCACCCCTAGGTTTCGGG 2520  
QY 2521 GCGTATATCTTAAGSCACATGCTATCGACCCCTAAACATCAGAACCCGGGTAAAGCACTC 2580  
Db 2521 GCGTATATCTTAAGSCACATGCTATCGACCCCTAAACATCAGAACCCGGGTAAAGCACTC 2580  
QY 2581 ACACGCGTGCCTCCCATACGTAATCTTACCTATGCAAGTTCCTTTCGCGACGGTGTTCG 2640  
Db 2581 ACACGCGTGCCTCCCATACGTAATCTTACCTATGCAAGTTCCTTTCGCGACGGTGTTCG 2640  
QY 2641 TCTGGGGCGCTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTCGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGCGCTGGAGCGGACTCGTCTG 2760  
Db 2701 ATCTCGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGCGCTGGAGCGGACTCGTCTG 2760  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACGTCGCACATCCAAACATCGAGGAGTG 2820  
Db 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACGTCGCACATCCAAACATCGAGGAGTG 2820  
QY 2821 GCTCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACCTCAATTTTCTGCCATTCGAAGAAGAAATGTGATGAGTTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACCTCAATTTTCTGCCATTCGAAGAAGAAATGTGATGAGTTCGCGCG 2940  
QY 2941 AAGCTCTCGGCTCGGACTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATTCGCTC 3000  
Db 2941 AAGCTCTCGGCTCGGACTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATTCGCTC 3000  
QY 3001 ATACCAACTAGCGGAGCGCTCATTTGCTGTAGCAACGAGCGCTCTAATGAGCGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGGAGCGCTCATTTGCTGTAGCAACGAGCGCTCTAATGAGCGGCTTTACC 3060  
QY 3061 GCGGATTTGCACTCAGTGTGATCGACTGCAATACATGTGTCCACCCAGACAGTTCGACTTCAGC 3120  
Db 3061 GCGGATTTGCACTCAGTGTGATCGACTGCAATACATGTGTCCACCCAGACAGTTCGACTTCAGC 3120  
QY 3121 CTGGACCGGACCTTACCAATTTGAGAGCAGCAACGTCGCCAAGACCGGCTGTACGCTCG 3180  
Db 3121 CTGGACCGGACCTTACCAATTTGAGAGCAGCAACGTCGCCAAGACCGGCTGTACGCTCG 3180

QY 3181 CAGCGGAGGACGACTGCTAGGGGACGAGTGGCAATTTACAGTGTGTGACTCCAGGA 3240  
Db |||||  
QY 3181 CAGCGGAGGACGAGACTGGTAGGGGACGAGTGGGCAATTTACAGTGTGTGACTCCAGGA 3240  
Db |||||  
QY 3241 GAACGGGCTCTGGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Db |||||  
QY 3241 GAACGGGCTCTGGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Db |||||  
QY 3301 GCTTGGTACGAGCTCAACCGCGCGAGACCTCAGTTAGTGTGCGGGCTTACTAAACACA 3360  
Db |||||  
QY 3301 GCTTGGTACGAGCTCAACCGCGCGAGACCTCAGTTAGTGTGCGGGCTTACTAAACACA 3360  
Db |||||  
QY 3361 CCAGGTTGCCCGTCTGCCAGACCATCTGGAGTCTGGGAGAGCGTCTTTACAGGCGCT 3420  
Db |||||  
QY 3361 CCAGGTTGCCCGTCTGCCAGACCATCTGGAGTCTGGGAGAGCGTCTTTACAGGCGCT 3420  
Db |||||  
QY 3421 ACCACATAGACGCCCAATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
Db |||||  
QY 3421 ACCACATAGACGCCCAATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
Db |||||  
QY 3481 CTGTAGCATACGAGCTACCGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATTCGTTGGAC 3540  
Db |||||  
QY 3481 CTGTAGCATACGAGCTACCGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATTCGTTGGAC 3540  
Db |||||  
QY 3541 CAATGTGGAGTGTCTCATACGCTAAAGCTAAGCTCAGCTGACGCGGCGCAACGCCCTGCTG 3600  
Db |||||  
QY 3541 CAATGTGGAGTGTCTCATACGCTAAAGCTAAGCTCAGCTGACGCGGCGCAACGCCCTGCTG 3600  
Db |||||  
QY 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAATACCAATATACATC 3660  
Db |||||  
QY 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAATACCAATATACATC 3660  
Db |||||  
QY 3661 ATGSCATGATGTGCGGTGACCTGGAGTGTCTACAGCACTGCTGGTGGAGGGA 3720  
Db |||||  
QY 3661 ATGSCATGATGTGCGGTGACCTGGAGTGTCTACAGCACTGCTGGTGGAGGGA 3720  
Db |||||  
QY 3721 GTCTAGCAGCTCTGGCGGCTATGCTGTACACAGCAGCGTGGTCAATTTGGGCGAG 3780  
Db |||||  
QY 3721 GTCTAGCAGCTCTGGCGGCTATGCTGTACACAGCAGCGTGGTCAATTTGGGCGAG 3780  
Db |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCTCATCTCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCTCATCTCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db |||||  
QY 3841 GATGAGATGGAAGTGGCGCTCACTCCTCTTACATCGAACAGGGAATGCAGCTCGCC 3900  
Db |||||  
QY 3841 GATGAGATGGAAGTGGCGCTCACTCCTCTTACATCGAACAGGGAATGCAGCTCGCC 3900  
Db |||||  
QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCAGCCACCAAGCAGCGAGCT 3960  
Db |||||  
QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCAGCCACCAAGCAGCGAGCT 3960  
Db |||||  
QY 3961 GCTGCTCCGTTGGTGAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
Db |||||  
QY 3961 GCTGCTCCGTTGGTGAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
Db |||||  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTGCTGGCAACCCC 4080  
Db |||||  
QY 4021 TGGAAATTCATACGCGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTGCTGGCAACCCC 4080  
Db |||||  
QY 4081 GCGATAGCATCACTGATGGGCAATTCACAGCTCTATACAGCCCGCTACACCCCAACAT 4140  
Db |||||  
QY 4081 GCGATAGCATCACTGATGGGCAATTCACAGCTCTATACAGCCCGCTACACCCCAACAT 4140  
Db |||||  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCGCAACTTGTCTCTCCAGCGCT 4200  
Db |||||  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCGCAACTTGTCTCTCCAGCGCT 4200  
Db |||||  
QY 4201 GCTTCTGCTTGTAGGCGCGGATCGCTGAGCGGCTGTGTGGCAGCATAGGCTTGGG 4260  
Db |||||  
QY 4201 GCTTCTGCTTGTAGGCGCGGATCGCTGAGCGGCTGTGTGGCAGCATAGGCTTGGG 4260  
Db |||||

QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCGCTCGTGCC 4320  
Db |||||  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCGCTCGTGCC 4320  
Db |||||  
QY 4321 TTTAAGGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
Db |||||  
QY 4321 TTTAAGGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
Db |||||  
QY 4381 ATCTCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGTCGAGCAGGATATCTCGTCCGAC 4440  
Db |||||  
QY 4381 ATCTCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGTCGAGCAGGATATCTCGTCCGAC 4440  
Db |||||  
QY 4441 GTGGGCGCAGGGAGGGGCTGTGAGTGTGATGAACCGGCTGATAGCCTTCTGCTTCGGG 4500  
Db |||||  
QY 4441 GTGGGCGCAGGGAGGGGCTGTGAGTGTGATGAACCGGCTGATAGCCTTCTGCTTCGGG 4500  
Db |||||  
QY 4501 GGTAAACACAGCTCTCCCCCAGCCTATGTGCTCAGAGCAGCTGACAGCAGTGTCACT 4560  
Db |||||  
QY 4501 GGTAAACACAGCTCTCCCCCAGCCTATGTGCTCAGAGCAGCTGACAGCAGTGTCACT 4560  
Db |||||  
QY 4561 CAGATCTCTCTAGTCTTTTACCATCACTCAGCTGTGAGAGGCTTCCACAGTGGATCAAC 4620  
Db |||||  
QY 4561 CAGATCTCTCTAGTCTTTTACCATCACTCAGCTGTGAGAGGCTTCCACAGTGGATCAAC 4620  
Db |||||  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAGAGATGTTGGGATTTGATATGC 4680  
Db |||||  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAGAGATGTTGGGATTTGATATGC 4680  
Db |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGTCCGCGGATTCGCGGA 4740  
Db |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCGCGGATTCGCGGGA 4740  
Db |||||  
QY 4741 GTCCCTCTCTCATGTCAACGTTGGGTACAAGGAGTCTGGCGGGCGCAGCGCATATG 4800  
Db |||||  
QY 4741 GTCCCTCTCTCATGTCAACGTTGGGTACAAGGAGTCTGGCGGGCGCAGCGCATATG 4800  
Db |||||  
QY 4801 CAACACACCTGCCCCATGTGAGGACAGATCACCGGACATGTGAAAAACGGTTCATGAGG 4860  
Db |||||  
QY 4801 CAACACACCTGCCCCATGTGAGGACAGATCACCGGACATGTGAAAAACGGTTCATGAGG 4860  
Db |||||  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACCGGTAC 4920  
Db |||||  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACCGGTAC 4920  
Db |||||  
QY 4921 ACCACGGGCCCCCTGACGCGCTCCCGCGCGCAAAATTAATCTAGGGGCTGTGGCGGGT 4980  
Db |||||  
QY 4921 ACCACGGGCCCCCTGACGCGCTCCCGCGCGCAAAATTAATCTAGGGGCTGTGGCGGGT 4980  
Db |||||  
QY 4981 GCTGCTGAGGATGAGTGGAGGTTACCGGGTGGGGATTTTCCACTAGTGAACGGGATG 5040  
Db |||||  
QY 4981 GCTGCTGAGGATGAGTGGAGGTTACCGGGTGGGGATTTTCCACTAGTGAACGGGATG 5040  
Db |||||  
QY 5041 ACCACTCACAAAGTAAAGTCCCGTGTCCGGCTCCCGCCCCCGAATTTCTTACAGAGATG 5100  
Db |||||  
QY 5041 ACCACTCACAAAGTAAAGTCCCGTGTCCGGCTCCCGCCCCCGAATTTCTTACAGAGATG 5100  
Db |||||  
QY 5101 GATGGGTGGGTTGACAGGTACGCTCCAGCGTGAACCCCTCTTACCGGAGGAGTTC 5160  
Db |||||  
QY 5101 GATGGGTGGGTTGACAGGTACGCTCCAGCGTGAACCCCTCTTACCGGAGGAGTTC 5160  
Db |||||  
QY 5161 ACATTCCTGTGCGGCTCAATCAATCACTGCTCAGCCCTCCCTTCCATCGGAGCGGAA 5220  
Db |||||  
QY 5161 ACATTCCTGTGCGGCTCAATCAATCACTGCTCAGCCCTCCCTTCCATCGGAGCGGAA 5220  
Db |||||  
QY 5221 CCGAGCTAGCAGTGTCTTCCATGCTCAGCCCTCCCTTCCATCGGAGCGGAA 5280  
Db |||||  
QY 5221 CCGAGCTAGCAGTGTCTTCCATGCTCAGCCCTCCCTTCCATCGGAGCGGAA 5280  
Db |||||  
QY 5281 GCTAAGGCTAGGCTGGCAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCAG 5340  
Db |||||  
QY 5281 GCTAAGGCTAGGCTGGCAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCAG 5340  
Db |||||  
QY 5341 CTGTCTGCGCTTCTTGAAGGCAACATGCACTACCGGCTCATGCTCCCGGAGCGTGCAC 5400  
Db |||||

Db	5341	CTGTCTGCGCCCTTCTTTGAAAGCAACATGCACTACCCGTTCATGACTCCCGCGACGCTGAC	5400	Db	6421	GTCCAAACAGAGAGAGGGGGCCGCAAGCCAGCTCGCTTTATCGTATTCOCAGATTGGGG	6480
QY	5401	CTCATCGAGCCCAACCTCTGTGGCGGCGAGAGATGGCGGGAACATACACCGCGTGGAG	5460	QY	6481	GTTCGTGTGTGCGAGAAATGCGCTTTACGATGTGGTCTTCCACCCCTCCCTCAGGCGGTG	6540
Db	5401	CTCATCGAGCCCAACCTCTGTGGCGGCGAGAGATGGCGGGAACATACACCGCGTGGAG	5460	Db	6481	GTTCGTGTGTGCGAGAAATGCGCTTTACGATGTGGTCTTCCACCCCTCCCTCAGGCGGTG	6540
QY	5461	TCGAAATAAGGTAGTAAATTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGGATGAG	5520	QY	6541	ATGGGCTCTTCATACGGAATTCGAATCTCTCTCGACAGCGGTGAGTTCCTGTGTGAAT	6600
Db	5461	TCGAAATAAGGTAGTAAATTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGATGAG	5520	Db	6541	ATGGGCTCTTCATACGGAATTCGAATCTCTCTCGACAGCGGTGAGTTCCTGTGTGAAT	6600
QY	5521	AGGGAAGTATCCGTTCCGGCGGAGATTCCTCGGAGGTCCAGGAAATTCCTCTCGAGCGATG	5580	QY	6601	GCTCGAAGCGAAGAAATGCGCTTATGGCTTCGCAATGACACCCGCTGTTTGTACTCA	6660
Db	5521	AGGGAAGTATCCGTTCCGGCGGAGATTCCTCGGAGGTCCAGGAAATTCCTCTCGAGCGATG	5580	Db	6601	GCTCGAAGCGAAGAAATGCGCTTATGGCTTCGCAATGACACCCGCTGTTTGTACTCA	6660
QY	5581	CCCATATGGGACCGCCCGGATTAACACCTTCACCTTTAGAGTCTTGGAAAGACCCGGAC	5640	QY	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
Db	5581	CCCATATGGGACCGCCCGGATTAACACCTTCACCTTTAGAGTCTTGGAAAGACCCGGAC	5640	Db	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
QY	5641	TACGTCCTCCAGTGTACACGGGTGTCCATTCGCCGCTGCCAAGCCCTCCGATACCA	5700	QY	6721	CCCGAAGCCAGACAGGCCATAAGGTTCGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG	6780
Db	5641	TACGTCCTCCAGTGTACACGGGTGTCCATTCGCCGCTGCCAAGCCCTCCGATACCA	5700	Db	6721	CCCGAAGCCAGACAGGCCATAAGGTTCGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG	6780
QY	5701	CCTCCAGGAGGAGGAGCGTGTCTGTCAAGATCTACCGTGTCTTCTGCTTGGCG	5760	QY	6781	ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTTCGCGCGAGCGGTGACTGACG	6840
Db	5701	CCTCCAGGAGGAGGAGCGTGTCTGTCAAGATCTACCGTGTCTTCTGCTTGGCG	5760	Db	6781	ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTTCGCGCGAGCGGTGACTGACG	6840
QY	5761	GAGCTCGCACAAGACCTTCGGCAGCTCCGAATCGTGGCGCTCGACAGCGCACGGCA	5820	QY	6841	ACCAGCTCGGCTAAATACCTCTCACATGTTTCTGAGGCGCGCTCGCGGCTGTGAGTGGC	6900
Db	5761	GAGCTCGCACAAGACCTTCGGCAGCTCCGAATCGTGGCGCTCGACAGCGCACGGCA	5820	Db	6841	ACCAGCTCGGCTAAATACCTCTCACATGTTTCTGAGGCGCGCTCGCGGCTGTGAGTGGC	6900
QY	5821	ACGGCTCTCTCAGCAGCCCTCCGACGCGGCGAGCGGGATCCGACGTTGAGTCTGTAC	5880	QY	6901	AAGCTCCAGGACTGCAAGTGTCTGATCGGAGACGACCTTGTCTGTATCTGTGAAAGC	6960
Db	5821	ACGGCTCTCTCAGCAGCCCTCCGACGCGGCGAGCGGGATCCGACGTTGAGTCTGTAC	5880	Db	6901	AAGCTCCAGGACTGCAAGTGTCTGATCGGAGACGACCTTGTCTGTATCTGTGAAAGC	6960
QY	5881	TCCTCCATGCCCCCTTGAAGGGGAGCCCGGGGATCCCGATCTCAGCGACGCGTCTTGG	5940	QY	6961	GCGGGACCCCAAGAGACGAGCGAGCTTACGCGGCTTACGAGGCTATGACTAGATAC	7020
Db	5881	TCCTCCATGCCCCCTTGAAGGGGAGCCCGGGGATCCCGATCTCAGCGACGCGTCTTGG	5940	Db	6961	GCGGGACCCCAAGAGACGAGCGAGCTTACGCGGCTTACGAGGCTATGACTAGATAC	7020
QY	5941	TTTACCGTAAAGAGAGGCTAGTAGGAGCGTGTCTGTCTGCTGCTGCTGCTGCTGCTG	6000	QY	7021	TCCTCCAAATGTGTAGTTCGCGCACGATGCTTGGCAAAAGGCTGATCAATCATGC	7080
Db	5941	TTTACCGTAAAGAGAGGCTAGTAGGAGCGTGTCTGTCTGCTGCTGCTGCTGCTGCTG	6000	Db	7021	TCCTCCAAATGTGTAGTTCGCGCACGATGCTTGGCAAAAGGCTGATCAATCATGC	7080
QY	6001	ACAGGCGCCCTGATCAGCCATCGCTGCGAGGAGAAACCAAGCTGCCATCAATGCACTG	6060	QY	7081	TCCTCCAAATGTGTAGTTCGCGCACGATGCTTGGCAAAAGGCTGATCAATCATGC	7140
Db	6001	ACAGGCGCCCTGATCAGCCATCGCTGCGAGGAGAAACCAAGCTGCCATCAATGCACTG	6060	Db	7081	TCCTCCAAATGTGTAGTTCGCGCACGATGCTTGGCAAAAGGCTGATCAATCATGC	7140
QY	6061	AGCAACTCTTTGCTCCGTCACCACTTGTGTATGCTACAACTCTGCTGCTGCTGCTGCTG	6120	QY	7141	GACCCACCAACCCCTTGGCGGCTGCGTGGAGACAGCTAGACACATCCAGTCAAT	7200
Db	6061	AGCAACTCTTTGCTCCGTCACCACTTGTGTATGCTACAACTCTGCTGCTGCTGCTGCTG	6120	Db	7141	GACCCACCAACCCCTTGGCGGCTGCGTGGAGACAGCTAGACACATCCAGTCAAT	7200
QY	6121	CTCGGGCAGAGAGGTCACCTTTGACAGACTGCGAGTCTGCGACGACCACTACCGGAC	6180	QY	7201	TCCTGGCTAGGCAACATCATCATGTATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
Db	6121	CTCGGGCAGAGAGGTCACCTTTGACAGACTGCGAGTCTGCGACGACCACTACCGGAC	6180	Db	7201	TCCTGGCTAGGCAACATCATCATGTATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
QY	6181	GTGCTCAAGGAGATGAAGGCGAGGCGTCCACAGTTAAGGCTAACTTCTATCCGTGGAG	6240	QY	7261	ACTCATTTCTTCTCCATCCCTTAGCTAGCTCAGGAAACAACTTGAAAAGCCCTAGATTGTGAG	7320
Db	6181	GTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTTAAGGCTAACTTCTATCCGTGGAG	6240	Db	7261	ACTCATTTCTTCTCCATCCCTTAGCTAGCTCAGGAAACAACTTGAAAAGCCCTAGATTGTGAG	7320
QY	6241	GAGGCTGTAAAGTGTAGCCGCCCACTTCGCGCAGATCTTAAATTTGGCTATGGGGCAAG	6300	QY	7321	ATCTACGGGCTGTTTACTCCATTGAGCCACTTGCCTACCTCAGATCAATCAACGACTC	7380
Db	6241	GAGGCTGTAAAGTGTAGCCGCCCACTTCGCGCAGATCTTAAATTTGGCTATGGGGCAAG	6300	Db	7321	ATCTACGGGCTGTTTACTCCATTGAGCCACTTGCCTACCTCAGATCAATCAACGACTC	7380
QY	6301	GAGCTCGGAACTTATCAGAGCGGCTTAAACACATCCGCTCCGTTGGAAGGACTTG	6360	QY	7381	CATGGCTTAGCCCACTTTTCACTCCATAGTTACTCTCCAGTGAATCAATAGGCTGGCT	7440
Db	6301	GAGCTCGGAACTTATCAGAGCGGCTTAAACACATCCGCTCCGTTGGAAGGACTTG	6360	Db	7381	CATGGCTTAGCCCACTTTTCACTCCATAGTTACTCTCCAGTGAATCAATAGGCTGGCT	7440
QY	6361	CTGGAAGACTGAGACCAATTTGACACACCATCATGCGCAAAATAGAGTTTCTGCG	6420	QY	7441	TCATGCTCAGGAAACTTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCCGGGCCAGAGT	7500
Db	6361	CTGGAAGACTGAGACCAATTTGACACACCATCATGCGCAAAATAGAGTTTCTGCG	6420	Db	7441	TCATGCTCAGGAAACTTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCCGGGCCAGAGT	7500
QY	6421	GTCCAAACAGAGAGGGGGCCGCAAGCCAGCTCGCTTTATCGTATTCOCAGATTGGGG	6480	QY	7501	GTCCGCGCTAGGCTACTGTCCCAAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC	7560
Db	6421	GTCCAAACAGAGAGGGGGCCGCAAGCCAGCTCGCTTTATCGTATTCOCAGATTGGGG	6480	Db	7501	GTCCGCGCTAGGCTACTGTCCCAAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC	7560

```
QY 7561 AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620
Db 7561 AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620
QY 7621 TTATCCAGCTGGTTCTGTTGCTGTTTACAGCGGGGAGACATATATACAGAGCTGTCCTGT 7680
Db 7621 TTATCCAGCTGGTTCTGTTGCTGTTTACAGCGGGGAGACATATATACAGAGCTGTCCTGT 7680
QY 7681 GCCCGACCCGCTGGTTCACTGTGGTCTACTCTCTAATTTCTGTAGGGGTAGGCACTAT 7740
Db 7681 GCCCGACCCGCTGGTTCACTGTGGTCTACTCTCTAATTTCTGTAGGGGTAGGCACTAT 7740
QY 7741 CTACTCTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
Db 7741 CTACTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7920
Db 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7920
QY 7921 TAGCTGTGAAGGTCGGTGAGCGCTTGACTGACGAGAGTCTGATAGTCTGCTCTCTG 7980
Db 7921 TAGCTGTGAAGGTCGGTGAGCGCTTGACTGACGAGAGTCTGATAGTCTGCTCTCTG 7980
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 5
US-10-125-920-1
; Sequence 1, Application US/10125920
; Publication No. US20030176503A1
; GENERAL INFORMATION:
; APPLICANT: Altamura, Sergio
; APPLICANT: Kock, Uwe
; TITLE OF INVENTION: TREATING HEPATITIS C VIRAL INFECTIONS
; TITLE OF INVENTION: WITH THIOSEMICARBAZONE COMPOUNDS
; FILE REFERENCE: IT0010Y
; CURRENT APPLICATION NUMBER: US/10/125.920
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,195
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10690
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid
US-10-125-920-1

Query Match 100.0%; Score 7987.4; DB 15; Length 10690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
Db 1 GCCAGCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
QY 61 TCITCAGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGCTGAGCTCCAGGAC 120
Db 61 TCITCAGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGCTGAGCTCCAGGAC 120
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCAG 180
```

```
QY 181 GACGACCGGTCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGGTGCCTCC 240
Db 181 GACGACCGGTCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGGTGCCTCC 240
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCCCTTGTGGTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCCCTTGTGGTACTGCTGATAGG 300
QY 301 GTGCTTGGAGTGCCTCCGGAGGTCCTGTAGACCGTGCACCATGAGCAAGATCCTAAAC 360
Db 301 GTGCTTGGAGTGCCTCCGGAGGTCCTGTAGACCGTGCACCATGAGCAAGATCCTAAAC 360
QY 361 CTCAAAGAAAACCAAGAGGCGCCCATGATTGAACAGATGGATTCACGAGGTTCTC 420
Db 361 CTCAAAGAAAACCAAGAGGCGCCCATGATTGAACAGATGGATTCACGAGGTTCTC 420
QY 421 CGGCGCTTGGGTGAGAGGCTATTCCGGCTATGACTGGGCAACAGACAAATCGGCTCT 480
Db 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAGACAAATCGGCTCT 480
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACG 540
Db 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACG 540
QY 541 ACCTGTCCGGTCCCTGAAATGAATCTGAGAGAGGAGCGCGGCTATCGTGGCTGCGCA 600
Db 541 ACCTGTCCGGTCCCTGAAATGAATCTGAGAGAGGAGCGCGGCTATCGTGGCTGCGCA 600
QY 601 CGACGGCGCTTCTTTCGCGAGCTGTGTCGACGTTGTCACCTGAAGCGGAGGAGCTGCG 660
Db 601 CGACGGCGCTTCTTTCGCGAGCTGTGTCGACGTTGTCACCTGAAGCGGAGGAGCTGCG 660
QY 661 TGCTATTGGGCGAGTGCCTGGGCGAGGATCTCTGTGTCATCTCACTTGTCTCTCCGAGA 720
Db 661 TGCTATTGGGCGAGTGCCTGGGCGAGGATCTCTGTGTCATCTCACTTGTCTCTCCGAGA 720
QY 721 AAGTATCCATAGCTGTGATGCAATGCGGCGGTGTCATAGCTTGTGATCCGGCTACCTGCC 780
Db 721 AAGTATCCATAGCTGTGATGCAATGCGGCGGTGTCATAGCTTGTGATCCGGCTACCTGCC 780
QY 781 CATTCGACCAACGAGCAATCGCATCGAGCGAGCACTGCTGATCGAGCGGAGCGGTC 840
Db 781 CATTCGACCAACGAGCAATCGCATCGAGCGAGCACTGCTGATCGAGCGGAGCGGTC 840
QY 841 TTGTGATCAGGATGATCTGAGAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTTCG 900
Db 841 TTGTGATCAGGATGATCTGAGAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTTCG 900
QY 901 CCAGGCTCAAGGCGCGCATGCCGACCGGAGGATCTCGTGTGACCCATGCGGATGCT 960
Db 901 CCAGGCTCAAGGCGCGCATGCCGACCGGAGGATCTCGTGTGACCCATGCGGATGCT 960
QY 961 GCTTGGCGAATATCATGTTGGAATGCGCGCTTTCTGGATTCATGACTGTGGCGGC 1020
Db 961 GCTTGGCGAATATCATGTTGGAATGCGCGCTTTCTGGATTCATGACTGTGGCGGC 1020
QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGC 1080
Db 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGC 1080
QY 1081 TTGGCGGCAATGGGCTGACCGCTTCTGTGTTTACGGTATCGCGCTCCCGATTCGC 1140
Db 1081 TTGGCGGCAATGGGCTGACCGCTTCTGTGTTTACGGTATCGCGCTCCCGATTCGC 1140
QY 1141 AGCGCATCGGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGTTAAACAGACCAAGC 1200
Db 1141 AGCGCATCGGCTTCTATCGCTTCTTGAAGAGTCTTCTGAGTTAAACAGACCAAGC 1200
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCTTACTTCTGTC 1260
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCTTACTTCTGTC 1260
```

QY 1261 CGAAGCCGCTTGGAAATAGGCGGTGTGGTTTGTCTATATGTTATTTTCCACCATATG 1320  
DB 1261 CGAAGCCGCTTGGAAATAGGCGGTGTGGTTTGTCTATATGTTATTTTCCACCATATG 1320  
QY 1321 CGGCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTTGACGAGCATTCCT 1380  
DB 1321 CGGCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTTGACGAGCATTCCT 1380  
QY 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTGTGAAGGAGCA 1440  
DB 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTGTGAAGGAGCA 1440  
QY 1441 GTTCTCTGGAAGCTCTTGAAGCAAAACAAAGTCTGTAGCACCTTTGCGAGGAGCG 1500  
DB 1441 GTTCTCTGGAAGCTCTTGAAGCAAAACAAAGTCTGTAGCACCTTTGCGAGGAGCG 1500  
QY 1501 AACCCGCCCACTGGGACAGGTGCTCTCGGSCCAAAGCCACGCTGATAAGATACACCT 1560  
DB 1501 AACCCGCCCACTGGGACAGGTGCTCTCGGSCCAAAGCCACGCTGATAAGATACACCT 1560  
QY 1561 GCAAGGCGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAGGCGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGATGCCCCAGAGGTACCCCATTTG 1680  
DB 1621 TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGATGCCCCAGAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAAA 1740  
QY 1741 AAGCTTAGGCCCCCGAACACCGGGAAGTGTGTTTCTTGAAGAAACAGATATAC 1800  
DB 1741 AAGCTTAGGCCCCCGAACACCGGGAAGTGTGTTTCTTGAAGAAACAGATATAC 1800  
QY 1801 ATGGCCCTATTACGCTTACTCCCAACAGCGAGGCTCTTGGCTGCATCATCACT 1860  
DB 1801 ATGGCCCTATTACGCTTACTCCCAACAGCGAGGCTCTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCACAGCGGGGACAGAACAGGTGAGGGGAGGTCCAGTGGTCTCCACGCA 1920  
DB 1861 AGCCTCACAGCGGGGACAGAACAGGTGAGGGGAGGTCCAGTGGTCTCCACGCA 1920  
QY 1921 ACACAATCTTTCTGGGCACTGGCTCAATGGCGTGTGGTGTGATCATGTGGC 1980  
DB 1921 ACACAATCTTTCTGGGCACTGGCTCAATGGCGTGTGGTGTGATCATGTGGC 1980  
QY 1981 GGCTCAAGACCTTGCAGCCCAAGGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
DB 1981 GGCTCAAGACCTTGCAGCCCAAGGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACCTGCTGGGTCAGGCGCCCGGCGGCGTTCTTTGACCAATGCACCTGC 2100  
DB 2041 CAGGACCTGCTGGGTCAGGCGCCCGGCGGCGTTCTTTGACCAATGCACCTGC 2100  
QY 2101 GGAGCTCGGACCTTTACTTGGTCAAGAGGATGCGGATGCTATCCGGTGCAGCGG 2160  
DB 2101 GGAGCTCGGACCTTTACTTGGTCAAGAGGATGCGGATGCTATCCGGTGCAGCGG 2160  
QY 2161 GGCGACAGAGGGGAGCCTTACTCTCCCGAGGCGGCTCTTACTTGAAGGCTCTTCG 2220  
DB 2161 GGCGACAGAGGGGAGCCTTACTCTCCCGAGGCGGCTCTTACTTGAAGGCTCTTCG 2220  
QY 2221 GGCGGTCCACTGTCTGCGCTCGGGCAAGCTGTGGGCAATCTTTGGGCTGCGGTGTC 2280  
DB 2221 GGCGGTCCACTGTCTGCGCTCGGGCAAGCTGTGGGCAATCTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCGAGGGGTGGAGGGGCTGGAATTTGTACCGCTGAGTCTATGGAACCACTATG 2340  
DB 2281 ACCGAGGGGTGGAGGGGCTGGAATTTGTACCGCTGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCAGGCAAACTCGTCCCGCTCGGCGGTACCGGACATTCAGGTG 2400

DB 2341 CGGTCCCGGCTTTCAGGCAAACTCGTCCCGCTCGGCGGTACCGGACATTCAGGTG 2400  
QY 2401 GCCCATCTACAGCCCTCTACTGTAGCGGCAAGACATAAGGTGCCGCTCGGTATGCA 2460  
DB 2401 GCCCATCTACAGCCCTCTACTGTAGCGGCAAGACATAAGGTGCCGCTCGGTATGCA 2460  
QY 2461 GCCAAGGGTATAAGGTGCTTCTGTAACCCGTGCTGCGCGCCACCTAGTGTTCGG 2520  
DB 2461 GCCAAGGGTATAAGGTGCTTCTGTAACCCGTGCTGCGCGCCACCTAGTGTTCGG 2520  
QY 2521 GGGTATATGCTTAAGGCAACATGGTATCGACCCCTAAACATCAGAAACCGGGTAAGGACATC 2580  
DB 2521 GGGTATATGCTTAAGGCAACATGGTATCGACCCCTAAACATCAGAAACCGGGTAAGGACATC 2580  
QY 2581 ACACGGGTGCCCCCATCAGGTACTCCACTATAGGCAAGTCTTTGCGGACGGTGTTC 2640  
DB 2581 ACACGGGTGCCCCCATCAGGTACTCCACTATAGGCAAGTCTTTGCGGACGGTGTTC 2640  
QY 2641 TCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB 2641 TCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCATCGGCACACTCTTGACATCATATATGATGATGAGTGCCACTCAACTG 2760  
DB 2701 ATCTGGGCATCGGCACACTCTTGACATCATATATGATGATGAGTGCCACTCGTGTG 2760  
QY 2761 CTGCGCACCGCTACCGCTCGGGATCGGTACCGTGCCACATCCAAACATCGAGGAGGT 2820  
DB 2761 CTGCGCACCGCTACCGCTCGGGATCGGTACCGTGCCACATCCAAACATCGAGGAGGT 2820  
QY 2821 GCTCTGTCAGCACTCGGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATC 2880  
DB 2821 GCTCTGTCAGCACTCGGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACCTCTATTTCTGCCATTCAGAGAAATGTGATGAGTTCGCGCG 2940  
DB 2881 AAGGGGGGAGGACCTCTATTTCTGCCATTCAGAGAAATGTGATGAGTTCGCGCG 2940  
QY 2941 AAGCTGTCGGCTCGGACTCAATGCTGTAGCATATTCGGGGCTTGTATGATCCGTC 3000  
DB 2941 AAGCTGTCGGCTCGGACTCAATGCTGTAGCATATTCGGGGCTTGTATGATCCGTC 3000  
QY 3001 ATACCAACTAGCGGACAGCTCATTTGTAGCAACGAGCGCTCTAATGACGGCTTACC 3060  
DB 3001 ATACCAACTAGCGGACAGCTCATTTGTAGCAACGAGCGCTCTAATGACGGCTTACC 3060  
QY 3061 GGCGATTTGACTCAGTGTATCGACTGCAATATGATGTGTCAACCCAGACAGTCCAGC 3120  
DB 3061 GGCGATTTGACTCAGTGTATCGACTGCAATATGATGTGTCAACCCAGACAGTCCAGC 3120  
QY 3121 CTGGACCGGACCTTACCATTTGAGACGACCGTGCACAAAGCGGCTTCAAGCTCG 3180  
DB 3121 CTGGACCGGACCTTACCATTTGAGACGACCGTGCACAAAGCGGCTTCAAGCTCG 3180  
QY 3181 CAGCGCGAGGAGGACTGTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
DB 3181 CAGCGCGAGGAGGACTGTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAAAGGCGCTTGGGCAATGTTGATTCCTGGTTCTGTGAGGTGCTATGACGGGCTGT 3300  
DB 3241 GAAAGGCGCTTGGGCAATGTTGATTCCTGGTTCTGTGAGGTGCTATGACGGGCTGT 3300  
QY 3301 GCTTGTACGAGCTACGCCCGGAGACCTCAGTTAGGTGCGGCTTACCTAAACACA 3360  
DB 3301 GCTTGTACGAGCTACGCCCGGAGACCTCAGTTAGGTGCGGCTTACCTAAACACA 3360  
QY 3361 CAGGGGTGCGGCTCTGCGGAGGACCTCTGAGTCTGAGGAGGCGTCTTTACAGGCTC 3420  
DB 3361 CAGGGGTGCGGCTCTGCGGAGGACCTCTGAGTCTGAGGAGGCGTCTTTACAGGCTC 3420  
QY 3421 ACCACATAGACGCCCATTTCTGTGCCAGCTAAGCAGGAGGACAACTTCCCTTAC 3480



Db 3421 ACCACATAGACGCCCAATTTCTTGTCACAGACTAAGCAGGACGAGACAACATTCCTCCCTAC 3480  
Qy 3481 CTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Qy 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCAACGCCCTGTGTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCAACGCCCTGTGTG 3600  
Qy 3601 TATAGGCTGGGAGCGGTTCAAAACAGAGTTACTACACACACCCCATAAACCAATACATC 3660  
Db 3601 TATAGGCTGGGAGCGGTTCAAAACAGAGTTACTACACACACCCCATAAACCAATACATC 3660  
Qy 3661 ATGGCATCATGTCCGGCTGACCTCGAGTTCGTCAGGACCTCGGTGTGTGTAGGCGGA 3720  
Db 3661 ATGGCATCATGTCCGGCTGACCTCGAGTTCGTCAGGACCTCGGTGTGTGTAGGCGGA 3720  
Qy 3721 GTCTAGCAGTCTGGCCGCGTATTGCTGACAAACAGGACGCTGCTCATTTGTGGGCAGG 3780  
Db 3721 GTCTAGCAGTCTGGCCGCGTATTGCTGACAAACAGGACGCTGCTCATTTGTGGGCAGG 3780  
Qy 3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCCGACAGGAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCCGACAGGAGTCTTTTACCGGAGTTC 3840  
Qy 3841 GATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGACAGGAAATCAGCTGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGACAGGAAATCAGCTGCC 3900  
Qy 3901 GAA CAATTCAAACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
Db 3901 GAA CAATTCAAACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
Qy 3961 GCTCTCCCGTGTGGAAATCAAGTGGGAGCCCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTCTCCCGTGTGGAAATCAAGTGGGAGCCCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Qy 4021 TGGAAATTCATCAGCGGATACAATATTTAGCAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAATATTTAGCAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
Qy 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACACCCCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACACCCCAACAT 4140  
Qy 4141 ACCCTCTGTTTACATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAATCTCTGGGGGATGGTGGCGGCCCAACTTGTCTTCCAGCGCT 4200  
Qy 4201 GCTTCTGCTTTCTAGGCGCGGCATCGCTGGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
Db 4201 GCTTCTGCTTTCTAGGCGCGGCATCGCTGGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
Qy 4261 AAGTGCTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGCGGCTCGTGCC 4320  
Db 4261 AAGTGCTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGCGGCTCGTGCC 4320  
Qy 4321 TTTAAGTGCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTCCCTGCT 4380  
Db 4321 TTTAAGTGCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTCCCTGCT 4380  
Qy 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGGGGTGTGGCCAGCGATCTGCTGGCCAC 4440  
Db 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGGGGTGTGGCCAGCGATCTGCTGGCCAC 4440  
Qy 4441 GTGGGCCAGGGAGGGGCTGTGSCAGTGAACCGGCTGTAGCGTTCTGCTTGGCG 4500  
Db 4441 GTGGGCCAGGGAGGGGCTGTGSCAGTGAACCGGCTGTAGCGTTCTGCTTGGCG 4500  
Qy 4501 GGTAAACACGCTCTCCCGCCACGACTATGTGCTGAGAGCAAGCTGAGCAAGCTGTCACT 4560  
Db 4501 GGTAAACACGCTCTCCCGCCACGACTATGTGCTGAGAGCAAGCTGAGCAAGCTGTCACT 4560

Qy 4561 CAGATCCTCTCTAGTCTTTACATCATCTAGCTGTGTAAGAGAGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTTACATCATCTAGCTGTGTAAGAGAGCTTCAACAGTGGATCAAC 4620  
Qy 4621 GAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGATGTTTGGATTTGATATGC 4680  
Db 4621 GAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGATGTTTGGATTTGATATGC 4680  
Qy 4681 ACGTGTGTGACTGATTTCAAGACCTTGGCTCCAGTCCAAAGCTCTCTCCCGCGATTGGCGGA 4740  
Db 4681 ACGTGTGTGACTGATTTCAAGACCTTGGCTCCAGTCCAAAGCTCTCTCCCGCGATTGGCGGA 4740  
Qy 4741 GTCCCTCTTCTCTCATGTCAAGTGGGTACAAAGAGTCTGGCGGGCGACGGCATCATG 4800  
Db 4741 GTCCCTCTTCTCTCATGTCAAGTGGGTACAAAGAGTCTGGCGGGCGACGGCATCATG 4800  
Qy 4801 CAACACCACTCCCATGTTGAGGACAGATCAACGAGCATGTGMAAAACGGTTCCATGAGG 4860  
Db 4801 CAACACCACTCCCATGTTGAGGACAGATCAACGAGCATGTGMAAAACGGTTCCATGAGG 4860  
Qy 4861 ATCGTGGGCTTAGGACCTGTAGTAACACGCTGGCATGGAACATTTCCCATTAACCGCTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAACACGCTGGCATGGAACATTTCCCATTAACCGCTAC 4920  
Qy 4921 ACCAGCGGCCCTGCAAGCCTTCCCGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG 4980  
Db 4921 ACCAGCGGCCCTGCAAGCCTTCCCGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG 4980  
Qy 4981 GCTCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTTCCACTAGTACGCGGATG 5040  
Db 4981 GCTCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTTCCACTAGTACGCGGATG 5040  
Qy 5041 ACCACTGACAACTGTAAGTGTCCGCTGTCAGGTTCCGGCCCCCAATTTCTCACAGAACTG 5100  
Db 5041 ACCACTGACAACTGTAAGTGTCCGCTGTCAGGTTCCGGCCCCCAATTTCTCACAGAACTG 5100  
Qy 5101 GATGGGTGGGTTGACAGTACGCTCCAGCGTGCAAAACCCCTCTACGCGGAGGCTC 5160  
Db 5101 GATGGGTGGGTTGACAGTACGCTCCAGCGTGCAAAACCCCTCTACGCGGAGGCTC 5160  
Qy 5161 ACATTCCTGTCGGCTCAATCAATACCTGTTGGGTTCAGAGTCCCATGCGAGCCGGA 5220  
Db 5161 ACATTCCTGTCGGCTCAATCAATACCTGTTGGGTTCAGAGTCCCATGCGAGCCGGA 5220  
Qy 5221 CCGGACCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGCGGAGACG 5280  
Db 5221 CCGGACCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGCGGAGACG 5280  
Qy 5281 GCTAAGCTAGGCTGGCGAGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Db 5281 GCTAAGCTAGGCTGGCGAGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Qy 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCGCTCATGACTCCCGGAGCTGAC 5400  
Db 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCGCTCATGACTCCCGGAGCTGAC 5400  
Qy 5401 CTCATCAGGCCAACCTCTCTGTCGGCGAGAGATGGCGGGCAACATCAACCGCTGGAG 5460  
Db 5401 CTCATCAGGCCAACCTCTCTGTCGGCGAGAGATGGCGGGCAACATCAACCGCTGGAG 5460  
Qy 5461 TCAGAAAATAAGTAGTATTTTGGACTCTTTTGGACTCTTTTGGAGGAGGATGAG 5520  
Db 5461 TCAGAAAATAAGTAGTATTTTGGACTCTTTTGGACTCTTTTGGAGGAGGATGAG 5520  
Qy 5521 AGGGAAGTATTCGTTCCGGCGGAGATCCTGCGGAGGTCCAGGAAATTTCCCTCAGCGATG 5580  
Db 5521 AGGGAAGTATTCGTTCCGGCGGAGATCCTGCGGAGGTCCAGGAAATTTCCCTCAGCGATG 5580  
Qy 5581 CCCATATGGGACGCGCGGATTAACACCTCTTCTAGAGTCTTCTAGAGGAGGACCCGGAC 5640  
Db 5581 CCCATATGGGACGCGCGGATTAACACCTCTTCTAGAGTCTTCTAGAGGAGGACCCGGAC 5640



QY	5641	TACGTCCCTCCAGTGGTACAGGGGTGCATTGCCGCTCCCAAGCGCCCTCCGATACCA	5700	Db	6721	CCCGAAGCCAGACAGGCCATAAGGTGCTTACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
Db	5641	TACGTCCCTCCAGTGGTACAGGGGTGCATTGCCGCTCCCAAGCGCCCTCCGATACCA	5700	QY	6781	ACTAATTTAAAGGCGAGAACTGCGGCTATCGCCGTCGCGCGAGCGGCTGCTACTGACG	6840
QY	5701	CCTCCACGAGGAAGAGGACGGTGTGCTGTCTGAGAACTTACCGTGTCTTCTGCTGGCG	5760	Db	6781	ACTAATTTAAAGGCGAGAACTGCGGCTATCGCCGTCGCGCGAGCGGCTGCTACTGACG	6840
Db	5701	CCTCCACGAGGAAGAGGACGGTGTGCTGTCTGAGAACTTACCGTGTCTTCTGCTGGCG	5760	QY	6841	ACCAGTCGGGTAAATACCTCATCTGTTACTTTGAAGGCGCTTCCGCGCTGTGAGCTGGG	6900
QY	5761	GAGTCGCGCAAAAGACCTTCGCGAGCTCCGAAATGTCGCGCGTTCGACAGCGGACGGCA	5820	Db	6841	ACCAGTCGGGTAAATACCTCATCTGTTACTTTGAAGGCGCTTCCGCGCTGTGAGCTGGG	6900
Db	5761	GAGTCGCGCAAAAGACCTTCGCGAGCTCCGAAATGTCGCGCGTTCGACAGCGGACGGCA	5820	QY	6901	AAGTCCAGGACTGACAGATGCTGTAATGCGAGACGACTTGTGCTGTTACTGTGTAAGGC	6960
QY	5821	AGGGCTCTCTGACAGAGCCTCCGACGCGGAGCGCGGATCCGACCTGAGTCGTAC	5880	Db	6901	AAGTCCAGGACTGACAGATGCTGTAATGCGAGACGACTTGTGCTGTTACTGTGTAAGGC	6960
Db	5821	AGGGCTCTCTGACAGAGCCTCCGACGCGGAGCGCGGATCCGACCTGAGTCGTAC	5880	QY	6961	GGGGGACCAAGAGGACGCGGAGCTTACCGGCTTCCAGGAGGCTATGACTAGATAC	7020
QY	5881	TCTTCATCCCCCTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGCGTCTTGG	5940	Db	6961	GGGGGACCAAGAGGACGCGGAGCTTACCGGCTTCCAGGAGGCTATGACTAGATAC	7020
Db	5881	TCTTCATCCCCCTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGCGTCTTGG	5940	QY	7021	TCTGCCCCCTTGGGACCCGCCCAACCCAGAATACGACTTGGAGTTGATAACATCATGC	7080
QY	5941	TCTACCGTAAGCGAGAGGCTAGTAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	Db	7021	TCTGCCCCCTTGGGACCCGCCCAACCCAGAATACGACTTGGAGTTGATAACATCATGC	7080
Db	5941	TCTACCGTAAGCGAGAGGCTAGTAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	QY	7081	TCTTCAATGTGTAGTCCGCGACGATGATCTGGCAAAAGGCTGTACTATCTCAACCGT	7140
QY	6001	ACAGGCGCCTGATCAGCGCATGCGCTGCGGAGAAACCAAGCTGCCATCAATGCACTG	6060	Db	7081	TCTTCAATGTGTAGTCCGCGACGATGATCTGGCAAAAGGCTGTACTATCTCAACCGT	7140
Db	6001	ACAGGCGCCTGATCAGCGCATGCGCTGCGGAGAAACCAAGCTGCCATCAATGCACTG	6060	QY	7141	GACCCACACCCCTTGGCGGCTGCTGGGACACAGCTAGACACACTCCAGTCAAT	7200
QY	6061	AGCAACTCTTGTGCTCGGTCAACCAACTTGGTCTATGCTACCAACTCTCGACGCAAGC	6120	Db	7141	GACCCACACCCCTTGGCGGCTGCTGGGACACAGCTAGACACACTCCAGTCAAT	7200
Db	6061	AGCAACTCTTGTGCTCGGTCAACCAACTTGGTCTATGCTACCAACTCTCGACGCAAGC	6120	QY	7201	TCTGTGCTAGGCAACATCATCATGATGCGCCACCTTGTGGCAAGGATGATCTGATG	7260
QY	6121	CTGCGGACAGAAAGGTACCTTTTGACAGACTCGAGTCTCGACGACCACTACCGGAC	6180	Db	7201	TCTGTGCTAGGCAACATCATCATGATGCGCCACCTTGTGGCAAGGATGATCTGATG	7260
Db	6121	CTGCGGACAGAAAGGTACCTTTTGACAGACTCGAGTCTCGACGACCACTACCGGAC	6180	QY	7261	ACTCATTTCTTCCATCTTCTAGCTCAGGAACAACTTGAAGGCTGTAGATTGTGAG	7320
QY	6181	GTGCTCAGGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAACTTCTATCCGTGGAG	6240	Db	7261	ACTCATTTCTTCCATCTTCTAGCTCAGGAACAACTTGAAGGCTGTAGATTGTGAG	7320
Db	6181	GTGCTCAGGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAACTTCTATCCGTGGAG	6240	QY	7321	ATCTAGCGGCTGTACTTCCATTGAGCCACTTACCTCAGATCATTTCAACGACTC	7380
QY	6241	GAAAGCTGTAAGCTGAGCGCCCAACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG	6300	Db	7321	ATCTAGCGGCTGTACTTCCATTGAGCCACTTACCTCAGATCATTTCAACGACTC	7380
Db	6241	GAAAGCTGTAAGCTGAGCGCCCAACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG	6300	QY	7381	CATGGCTTAGCGCATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT	7440
QY	6301	GAGCTCGGAAACCTATCCAGCAAGCGGTAAACCACTCCGCTCCGCTGCGAAGGACTTG	6360	Db	7381	CATGGCTTAGCGCATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT	7440
Db	6301	GAGCTCGGAAACCTATCCAGCAAGCGGTAAACCACTCCGCTCCGCTGCGAAGGACTTG	6360	QY	7441	TCATGCTCAGGAAAATTGGGGTACCGCTTCCGAGTCTGGAGACATCGGCGCAGAAGT	7500
QY	6361	CTGGAAGACTGAGACACCAATTGACACCACTATGACCACTATGCGAATAATGAGTTTCTGC	6420	Db	7441	TCATGCTCAGGAAAATTGGGGTACCGCTTCCGAGTCTGGAGACATCGGCGCAGAAGT	7500
Db	6361	CTGGAAGACTGAGACACCAATTGACACCACTATGACCACTATGCGAATAATGAGTTTCTGC	6420	QY	7501	GTCCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGCAAGTACTCTTTC	7560
QY	6421	GTCCAAACAGAGAGGGGGCGCAAGCAGCTGCGCTTATCGTATTTCCAGATTTGGG	6480	Db	7501	GTCCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGCAAGTACTCTTTC	7560
Db	6421	GTCCAAACAGAGAGGGGGCGCAAGCAGCTGCGCTTATCGTATTTCCAGATTTGGG	6480	QY	7561	AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCATCCCGCTGGTCCAGTTGGAT	7620
QY	6481	GTTCGTGTGCGAGAAAATGGCCCTTACGATGTGTCTCCACCTCCCTCAGGCGCTG	6540	Db	7561	AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCATCCCGCTGGTCCAGTTGGAT	7620
Db	6481	GTTCGTGTGCGAGAAAATGGCCCTTACGATGTGTCTCCACCTCCCTCAGGCGCTG	6540	QY	7621	TTATCCAGCTGGTGTGTTAGCGGGGAGACATATATCAAGCTGTCTCGT	7680
QY	6541	ATGGGCTCTTACACGATTCGAATCTCTCGGACAGGGGTCCGAGTTCTCTGTTGAAT	6600	Db	7621	TTATCCAGCTGGTGTGTTAGCGGGGAGACATATATCAAGCTGTCTCGT	7680
Db	6541	ATGGGCTCTTACACGATTCGAATCTCTCGGACAGGGGTCCGAGTTCTCTGTTGAAT	6600	QY	7681	GCCCGACCCGCTGTGTTCTGCTACTTCTTCTAGGGGTAGGATCATCTAT	7740
QY	6601	GCCTGGAAGCGAAGAAATGCCCTATGGGCTTGCATATGACACCCGCTGTTTGA	6660	Db	7681	GCCCGACCCGCTGTGTTCTGCTACTTCTTCTAGGGGTAGGATCATCTAT	7740
Db	6601	GCCTGGAAGCGAAGAAATGCCCTATGGGCTTGCATATGACACCCGCTGTTTGA	6660	QY	7741	CTACTCCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT	7800
QY	6661	ACGGTCACTGAGATGATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGCC	6720	Db	7741	CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT	7800
Db	6661	ACGGTCACTGAGATGATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGCC	6720	QY	7801	TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	7860
QY	6721	CCCGAAGCCAGACAGGCCATAAGGTGCTTACAGAGCGGCTTTTACATCGGGGGCCCCCTG	6780				



Db 1501 AACCCCCCAGCTGGGACAGGTGCTCTGGGGCAGAAACCCACGCTGTATATAGATACACT 1560  
 QY 1561 GCAAGGGGACAAACCCAGTGCACAGTGTGGATAGTTGTGGAAAGAGTCAAA 1620  
 Db 1561 GCAAGGGGACAAACCCAGTGCACAGTGTGGATAGTTGTGGAAAGAGTCAAA 1620  
 QY 1621 TGGCTCTCTTAAGGGTATTCAAGGGGCTGAAGGATGCCAGAAAGTACCCATTGT 1680  
 Db 1621 TGGCTCTCTTAAGGGTATTCAAGGGGCTGAAGGATGCCAGAAAGTACCCATTGT 1680  
 QY 1681 ATGGGATCTGAATCTGGGCTCGGTGCACATCTTTTACATGTGTAGTCGAGGTTAAAA 1740  
 Db 1681 ATGGGATCTGAATCTGGGCTCGGTGCACATCTTTTACATGTGTAGTCGAGGTTAAAA 1740  
 QY 1741 AACGCTTAGGCCCCCGAACACCGGGACGCTGGTTTCTTTGAAAAACACGATAATACC 1800  
 Db 1741 AACGCTTAGGCCCCCGAACACCGGGACGCTGGTTTCTTTGAAAAACACGATAATACC 1800  
 QY 1801 ATGGGCTTATTAAGGCTTACTCCCAACAGAGCGAGGCTACTTTGGCTGCATCATCACT 1860  
 Db 1801 ATGGGCTTATTAAGGCTTACTCCCAACAGAGCGAGGCTACTTTGGCTGCATCATCACT 1860  
 QY 1861 AGCCTCACAGGCGGACAGGAAACAGAGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
 Db 1861 AGCCTCACAGGCGGACAGGAAACAGAGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
 QY 1921 ACACAACTTTCTGGCGACCTTGGCTCAATGGCGTGTGGACTGTCTATCATGGTGCC 1980  
 Db 1921 ACACAACTTTCTGGCGACCTTGGCTCAATGGCGTGTGGACTGTCTATCATGGTGCC 1980  
 QY 1981 GGTCAAGACCTTTCGGGCTTACTCCCAACAGAGCGAGGCTACTTTGGCTGCATCATCACT 2040  
 Db 1981 GGTCAAGACCTTTCGGGCTTACTCCCAACAGAGCGAGGCTACTTTGGCTGCATCATCACT 2040  
 QY 2041 CAGGACCTGTGGCTGGCAAGCGCCCCCGGGGCGCTTCTTTGACACCATGCACTGC 2100  
 Db 2041 CAGGACCTGTGGCTGGCAAGCGCCCCCGGGGCGCTTCTTTGACACCATGCACTGC 2100  
 QY 2101 GGCAGCTCGGACCTTTACTTGTGTACAGGCAATGCGATGTCAATCCGGTGGCGCGGG 2160  
 Db 2101 GGCAGCTCGGACCTTTACTTGTGTACAGGCAATGCGATGTCAATCCGGTGGCGCGGG 2160  
 QY 2161 GGCAGCAGAGGGGAGCTTCTCCCGGAGCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
 Db 2161 GGCAGCAGAGGGGAGCTTCTCTCCCGGAGCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
 QY 2221 GCGGCTCCACTGCTCTGCGGCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGCTGTGC 2280  
 Db 2221 GCGGCTCCACTGCTCTGCGGCTCGGGGACGCTGTGGGCACTTTTGGGCTGCGGCTGTGC 2280  
 QY 2281 ACCGAGGGGTTGCGAAGCGGTGGAATTTGTATACCGTTCAGTCTATGGAAACCACTATG 2340  
 Db 2281 ACCGAGGGGTTGCGAAGCGGTGGAATTTGTATACCGTTCAGTCTATGGAAACCACTATG 2340  
 QY 2341 CGGTCCCGCTTCTCAGCACTCTGCTCCCTCCCGGCTTACCGGACATTTCCAGGTG 2400  
 Db 2341 CGGTCCCGCTTCTCAGCACTCTGCTCCCTCCCGGCTTACCGGACATTTCCAGGTG 2400  
 QY 2401 GGCATCTACAGCGCTTCTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCA 2460  
 Db 2401 GGCATCTACAGCGCTTCTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCA 2460  
 QY 2461 GGCAGGGGTATAGGTGCTTCTGAAACCGCTCGGTGCGGCGGCTAGGTTTCGGG 2520  
 Db 2461 GGCAGGGGTATAGGTGCTTCTGAAACCGCTCGGTGCGGCGGCTAGGTTTCGGG 2520  
 QY 2521 GCGTATATGCTAAGGACACATGATGACCCCTTAACATGAGAACCGGGTAAAGACCATC 2580  
 Db 2521 GCGTATATGCTAAGGACACATGATGACCCCTTAACATGAGAACCGGGTAAAGACCATC 2580  
 QY 2581 ACCAGGGGTGCCCCCATCACTACTCCACTATGGCAAGTTTCTTTCGCGAGCGGTGTGC 2640  
 Db 2581 ACCAGGGGTGCCCCCATCACTACTCCACTATGGCAAGTTTCTTTCGCGAGCGGTGTGC 2640

QY 2641 TCTGGGGCGCTATGACATCATATATGTATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
 Db 2641 TCTGGGGCGCTATGACATCATATATGTATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
 QY 2701 ATCTGGGCATCGGCACAGTCTCTGCAACCAAGCGGAGACGGCTGGAGCGCACTCGTCGTC 2760  
 Db 2701 ATCTGGGCATCGGCACAGTCTCTGCAACCAAGCGGAGACGGCTGGAGCGCACTCGTCGTC 2760  
 QY 2761 CTGCGCACCGCTACCGCTTCGGGATCGGTCCACCGTGCACATCCAAACATCCAGAGGTC 2820  
 Db 2761 CTGCGCACCGCTACCGCTTCGGGATCGGTCCACCGTGCACATCCAAACATCCAGAGGTC 2820  
 QY 2821 GCTCTGTCCAGCACCTGAGAAATCCCTTTTATGCAAGAGCCATCCCATTCAGAGCCATC 2880  
 Db 2821 GCTCTGTCCAGCACCTGAGAAATCCCTTTTATGCAAGAGCCATCCCATTCAGAGCCATC 2880  
 QY 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCACAAAGAAATGTATGATGTCGCGG 2940  
 Db 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCACAAAGAAATGTATGATGTCGCGG 2940  
 QY 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGATCCGTC 3000  
 Db 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGATCCGTC 3000  
 QY 3001 ATACCAACTAGCGGAGAGCTCAATGCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
 Db 3001 ATACCAACTAGCGGAGAGCTCAATGCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
 QY 3061 GCGCATTTTCCGATCTAGTGTACGATGCAATATGTCACCCAGACAGTGTGACTCCAGCA 3120  
 Db 3061 GCGCATTTTCCGATCTAGTGTACGATGCAATATGTCACCCAGACAGTGTGACTCCAGCA 3120  
 QY 3121 CTGAGCCGACCTTACATTTAGAGACGACGCTGTCACAGACGCGGTGTGACTCCAGCA 3180  
 Db 3121 CTGAGCCGACCTTACATTTAGAGACGACGCTGTCACAGACGCGGTGTGACTCCAGCA 3180  
 QY 3181 CAGCGCGAGGAGGACTGTTAGGGGAGGATGGGCTTTACAGGTTTGTGACTCCAGCA 3240  
 Db 3181 CAGCGCGAGGAGGACTGTTAGGGGAGGATGGGCTTTACAGGTTTGTGACTCCAGCA 3240  
 QY 3241 GAACGGCTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
 Db 3241 GAACGGCTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
 QY 3301 GCTTGTAGAGCTACGCGGCGGAGACCTCAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 3360  
 Db 3301 GCTTGTAGAGCTACGCGGCGGAGACCTCAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 3360  
 QY 3361 CCAAGGTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
 Db 3361 CCAAGGTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
 QY 3421 ACCACATAGAGCCCATTTTGTCCAGACTAAGAGGAGAGAGCAATCCCTGCTAC 3480  
 Db 3421 ACCACATAGAGCCCATTTTGTCCAGACTAAGAGGAGAGAGCAATCCCTGCTAC 3480  
 QY 3481 CTGTTAGCATACAGGCTACGCTGCTGCGCAGGCTCAGCTCCACTCCATCGTGGGAC 3540  
 Db 3481 CTGTTAGCATACAGGCTACGCTGCTGCGCAGGCTCAGCTCCACTCCATCGTGGGAC 3540  
 QY 3541 CAAATGTGAAGTGTCTCATACGCTTAAAGCTTCTGCTGCAACGGGCAACCGCCCTGCTG 3600  
 Db 3541 CAAATGTGAAGTGTCTCATACGCTTAAAGCTTCTGCTGCAACGGGCAACCGCCCTGCTG 3600  
 QY 3601 TATAGGCTGGAGCGGCTTCAAAACGAGGTTTACTACACACACCCCATTAACCAATATATC 3660  
 Db 3601 TATAGGCTGGAGCGGCTTCAAAACGAGGTTTACTACACACACCCCATTAACCAATATATC 3660  
 QY 3661 ATGGCATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720  
 Db 3661 ATGGCATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720

QY 3721 GTCTAGCAGCTCTGCGCGCGTATTGCTGCAAAACAGGCGAGTGGTCTATTGTGGCAGG 3780  
Db 3721 GTCTAGCAGCTCTGCGCGCGTATTGCTGCAAAACAGGCGAGTGGTCTATTGTGGCAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCGCGCATCATTCGCGACAGGAAAGTCTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGCGCATCATTCGCGACAGGAAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGGGCGCTCACACCTCCCTTACATCGAACAGGAGTGCAGTCGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGGCGCTCACACCTCCCTTACATCGAACAGGAGTGCAGTCGCC 3900  
QY 3901 GAACAATTCAACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGCAGGCT 3960  
Db 3901 GAACAATTCAACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGGAGGCT 3960  
QY 3961 GCTGCTCCGCTGGTGGAAATCAAGTGGCGGACCTTCGAAGCCCTTCGCGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGGAAATCAAGTGGCGGACCTTCGAAGCCCTTCGCGCGAAGCATATG 4020  
QY 4021 TGGAAATTTATCAGCGGAGTACAAATATTAGCAGGCTTTGTCACCTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTTATCAGCGGAGTACAAATATTAGCAGGCTTTGTCACCTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACCGCCGCTCACCAACCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACCGCCGCTCACCAACCACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTGTAAGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
Db 4201 GCTTCTGCTTTGTAAGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
QY 4261 AAGGTGCTTGTGATATTTTGGCAGTTATGAGCAGGCGTGGCAGGCGCTCGTGSCC 4320  
Db 4261 AAGGTGCTTGTGATATTTTGGCAGTTATGAGCAGGCGTGGCAGGCGCTCGTGSCC 4320  
QY 4321 TTTAAGGTGATAGCGGGAGATGCTCTTCAACGAGACCTGGTTAACTACTCCCTGCT 4380  
Db 4321 TTTAAGGTGATAGCGGGAGATGCTCTTCAACGAGACCTGGTTAACTACTCCCTGCT 4380  
QY 4381 ATCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGCGATCTGCTGCGCAC 4440  
Db 4381 ATCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGCGATCTGCTGCGCAC 4440  
QY 4441 GTGGGCCCCAGGGAGGGGCTGTGAGTGAATGAACCGGCTGTAGCGTTGCTTCCGGG 4500  
Db 4441 GTGGGCCCCAGGGAGGGGCTGTGAGTGAATGAACCGGCTGTAGCGTTGCTTCCGGG 4500  
QY 4501 GGTAAACACGCTCCCCCAGCATATGTGCTGAGAGCGACGTGAGCAGCGTCACT 4560  
Db 4501 GGTAAACACGCTCCCCCAGCATATGTGCTGAGAGCGACGTGAGCAGCGTCACT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGTCTGAGAGGCTTCCACAGTGAATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGTCTGAGAGGCTTCCACAGTGAATCAAC 4620  
QY 4621 GAGGATGTCACCGCATCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Db 4621 GAGGATGTCACCGCATCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
QY 4681 ACGGTGTTGATGATTTCAAGACTGCTGCTCAAGTCTTCTGCGCGGATTTGCCGGA 4740  
Db 4681 ACGGTGTTGATGATTTCAAGACTGCTGCTCAAGTCTTCTGCGCGGATTTGCCGGA 4740  
QY 4741 GTCCCCCTTCTTCAATGTCAGTGGGTACAGGAGTCTGGCGGCGCAGCGATCATG 4800  
Db 4741 GTCCCCCTTCTTCAATGTCAGTGGGTACAGGAGTCTGGCGGCGCAGCGATCATG 4800  
QY 4801 CAAACCACTGCCCCATGTGGAGCAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860

Db 4801 CAAACCACTGCCCCATGTGGAGCAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860  
QY 4861 ATCTGTTGGGCTTAGACCTCTAGTAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
Db 4861 ATCTGTTGGGCTTAGACCTCTAGTAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
QY 4921 ACCACGGGCTTGCACGCGCTTCCCGGCGCAAAATTTCTTAGGGCGCTGTGGGGGTG 4980  
Db 4921 ACCACGGGCTTGCACGCGCTTCCCGGCGCAAAATTTCTTAGGGCGCTGTGGGGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGTTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
QY 5041 ACCACTGACAAAGTAAAGTGGCCCTGTCAGGTTCCGGGCCCGCAATTTCTTACACAAAGT 5100  
Db 5041 ACCACTGACAAAGTAAAGTGGCCCTGTCAGGTTCCGGGCCCGCAATTTCTTACACAAAGT 5100  
QY 5101 GATGGGTCGCGTTGCAACAGTACGCTCAGCGTGCAAAACCCCTCTTACGGGAGGAGTTC 5160  
Db 5101 GATGGGTCGCGTTGCAACAGTACGCTCAGCGTGCAAAACCCCTCTTACGGGAGGAGTTC 5160  
QY 5161 ACATTTCTGTCGCGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTTCTGTCGCGCTCAATCAATACCTGTTGGTTCACAGCTCCCATGCGAGCCGAA 5220  
QY 5221 CCGACGTAGCAGTCTCATTCCATGCTCACCGACCCCTCCCATTTACCGCGGAGAG 5280  
Db 5221 CCGACGTAGCAGTCTCATTCCATGCTCACCGACCCCTCCCATTTACCGCGGAGAG 5280  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCTCTCTTGGCCAGTCTATAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCTCTCTTGGCCAGTCTATAGCTAGCCAG 5340  
QY 5341 CTGCTCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGATCCCGGAGCGCTGAC 5400  
Db 5341 CTGCTCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGATCCCGGAGCGCTGAC 5400  
QY 5401 CTATCAGAGCCAACTCTCTGTCGCGCAGGAGATGGCGGAAACATCACCCGCTGGAG 5460  
Db 5401 CTATCAGAGCCAACTCTCTGTCGCGCAGGAGATGGCGGAAACATCACCCGCTGGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTTCCAAAGCGGAGGATGAG 5520  
Db 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTTCCAAAGCGGAGGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCGGAGGTCAGGAAATTTCCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCGGAGGTCAGGAAATTTCCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGACGCGCCGATTAACACCTCTCAGTCTTAGAGTCTCGGAGGACCCGAC 5640  
Db 5581 CCCATATGGGACGCGCCGATTAACACCTCTCAGTCTTAGAGTCTCGGAGGACCCGAC 5640  
QY 5641 TAGCTCCTCCAGTGTGTAACGGGTGTCAATGCGCTTCCCAAGGCGCTTCCGATACCA 5700  
Db 5641 TAGCTCCTCCAGTGTGTAACGGGTGTCAATGCGCTTCCCAAGGCGCTTCCGATACCA 5700  
QY 5701 CTTCCACGGAGGAGGAGGCTTGTCTGTCTAGATCTACCGTGTCTTCTGCTGGCG 5760  
Db 5701 CTTCCACGGAGGAGGAGGCTTGTCTGTCTAGATCTACCGTGTCTTCTGCTGGCG 5760  
QY 5761 GAGCTGCCACAAAGACCTTTCCGAGCTCCGAATCTCGSCGCTCGACAGCGCAGCGCA 5820  
Db 5761 GAGCTGCCACAAAGACCTTTCCGAGCTCCGAATCTCGSCGCTCGACAGCGCAGCGCA 5820  
QY 5821 ACGGCTCTCTGACCCAGCCCTCCGACGAGCGGCGAGCGGGATCCGACGTTGAGTCTG 5880  
Db 5821 ACGGCTCTCTGACCCAGCCCTCCGACGAGCGGCGAGCGGGATCCGACGTTGAGTCTG 5880  
QY 5881 TCTCTCATGCCCCCTTTCAGGGGAGCGGGGATCCCGATCTCAGCGAGCGGCTTTGG 5940



Publication No. US2002015133A1  
; GENERAL INFORMATION:  
; APPLICANT: ANADYS Pharmaceuticals, Inc.  
; APPLICANT: Bichko, Vadim  
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION  
; FILE REFERENCE: 0342/1H3950S1  
; CURRENT APPLICATION NUMBER: US/10/005,469  
; CURRENT FILING DATE: 2002-04-18  
; PRIOR APPLICATION NUMBER: US 60/245,866  
; PRIOR FILING DATE: 2000-11-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 7992  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HCV Replicon from cell line HCVr24  
US-10-005-469-6

Query Match 100.0%; Score 7985.8; DB 13; Length 7992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGAACTACTG 60  
DB 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGAACTACTG 60

QY 61 TCITCAGCAGAAAGCGTCTAGCAGATGCGCGTTAGTATGAGTGTGCGTGAGCCTCCAGGAC 120  
DB 61 TCITCAGCAGAAAGCGTCTAGCAGATGCGCGTTAGTATGAGTGTGCGTGAGCCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCTAGTGTCTCGGACCGGTGAGTACACCGGAATTCGAC 180  
DB 121 CCCCCCTCCCGGAGAGCCTAGTGTCTCGGACCGGTGAGTACACCGGAATTCGAC 180

QY 181 GACGACCGGTCCTTTCTTGGATCAACCCCGTCAATGCTGGAGATTGGCGTGCCTCC 240  
DB 181 GACGACCGGTCCTTTCTTGGATCAACCCCGTCAATGCTGGAGATTGGCGTGCCTCC 240

QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGACTGCTGATAGG 300  
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGACTGCTGATAGG 300

QY 301 GTGCTTCGAGTGCCTCCGCGAGTCTGTAGACCGGTGACCATGACGACGATCCTAAAC 360  
DB 301 GTGCTTCGAGTGCCTCCGCGAGTCTGTAGACCGGTGACCATGACGACGATCCTAAAC 360

QY 361 CTCAAGAAAAACAAAGGCGCGCCATGATTGAACAAGATGGATTGCACGAGTTCTC 420  
DB 361 CTCAAGAAAAACAAAGGCGCGCCATGATTGAACAAGATGGATTGCACGAGTTCTC 420

QY 421 CGGCGCTTGGTGGAGAGCTATTGGCTATGACTGGGACAAACAGCAATCGGCTGCT 480  
DB 421 CGGCGCTTGGTGGAGAGCTATTGGCTATGACTGGGACAAACAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGCTGTTCCGCGTGTACGCGAGGCGCGCGCTTCTTTTGTCAAGACCG 540  
DB 481 CTGATGCGCGCTGTTCCGCGTGTACGCGAGGCGCGCGCTTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGCTGCCCTGAATGAATGACGAGCAGGCGAGCGCGCTATCGTGGCTGSCCA 600  
DB 541 ACCTGTCCGCTGCCCTGAATGAATGACGAGCAGGCGAGCGCGCTATCGTGGCTGSCCA 600

QY 601 CGAGGCGCTTCTTGGCAGTGTGTCTGACGTTGTCACTGAAGCGGGGAAGGACTGGC 660  
DB 601 CGAGGCGCTTCTTGGCAGTGTGTCTGACGTTGTCACTGAAGCGGGGAAGGACTGGC 660

QY 661 TGTATTGGCGAAGTCCCGGCGAGATCTCTGTATCTCACTTGTCTCTGCGGAGA 720  
DB 661 TGTATTGGCGAAGTCCCGGCGAGATCTCTGTATCTCACTTGTCTCTGCGGAGA 720

QY 721 AAGTATCCATCATGGCTGATGCAATGCGGCGGTGCATACGCTTGATCCGGCTACCTGCC 780

DB 721 AAGTATCCATCATGGCTGATGCAATGCGGCGGTGCATACGCTTGCATCGGTACTCGCTGCC 780

QY 781 CATTCGACCAACCAAGCAAAATCCCATCGAGCGAGCAGTACTCGATGGAAGCGGTC 840

DB 781 CATTCGACCAACCAAGCAAAATCCCATCGAGCGAGCAGTACTCGATGGAAGCGGTC 840

QY 841 TTGTGATCAGATGATCTGACGAGAGCATCAGGGGCTCGCGCCAGCGAATGTTGCG 900

DB 841 TTGTGATCAGATGATCTGACGAGAGCATCAGGGGCTCGCGCCAGCGAATGTTGCG 900

QY 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTGATCCCATGCGGATGCT 960

DB 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTGATCCCATGCGGATGCT 960

QY 961 GCTTGCAGATATCATGGTGAATAATGCCGCTTTCTGGATTTCATCGACTGTGCCGCG 1020

DB 961 GCTTGCAGATATCATGGTGAATAATGCCGCTTTCTGGATTTCATCGACTGTGCCGCG 1020

QY 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080

DB 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080

QY 1081 TTGGCGGAGATGGCTGACCGCTTCTCTGCTTACGGTATCGCCGCTCCGATTCGC 1140

DB 1081 TTGGCGGAGATGGCTGACCGCTTCTCTGCTTACGGTATCGCCGCTCCGATTCGC 1140

QY 1141 AGCGATCGCTTCTATCGCCTTCTTGACGAGTCTCTCTGAGTTTAAACAGACCAACAG 1200

DB 1141 AGCGATCGCTTCTATCGCCTTCTTGACGAGTCTCTCTGAGTTTAAACAGACCAACAG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAACTACTGCG 1260

DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAACTACTGCG 1260

QY 1261 CGAAGCGCTTGAATTAAGCGGCTGCTGCTTCTATATGTTTATTTCCACCATATG 1320

DB 1261 CGAAGCGCTTGAATTAAGCGGCTGCTGCTTCTATATGTTTATTTCCACCATATG 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTGAAGCATTCCT 1380

DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTGAAGCATTCCT 1380

QY 1381 AGGGTCTTTTCCCTCTCGCAAGGAATCAAGGTCTGTGATGTCGTAAGGAAGCA 1440

DB 1381 AGGGTCTTTTCCCTCTCGCAAGGAATCAAGGTCTGTGATGTCGTAAGGAAGCA 1440

QY 1441 GTTCTCTGGAAGCTTCTTGAAGCAACAAACGCTGTGAGGACCTTTGAGGACGGG 1500

DB 1441 GTTCTCTGGAAGCTTCTTGAAGCAACAAACGCTGTGAGGACCTTTGAGGACGGG 1500

QY 1501 AACCCCGCCTCGGACAGGTGCTCTGGCGCAAAAGCCACGTGTATAGATACACCT 1560

DB 1501 AACCCCGCCTCGGACAGGTGCTCTGGCGCAAAAGCCACGTGTATAGATACACCT 1560

QY 1561 GMAAGCGGCAACACCCAGTCCAGTGTGAGTGTGATGTTGTAAGAGTCAAA 1620

DB 1561 GMAAGCGGCAACACCCAGTCCAGTGTGAGTGTGATGTTGTAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCTATTCACCAAGGCGCTGAAGATGCCAGAGGTACCCCATTTG 1680

DB 1621 TGGCTCTCTCAAGCTATTCACCAAGGCGCTGAAGATGCCAGAGGTACCCCATTTG 1680

QY 1681 ATGGATCTGATCTGGGCTCGTGACATGTTTACATGCTTGTAGTGCAGGTTAAA 1740

DB 1681 ATGGATCTGATCTGGGCTCGTGACATGTTTACATGCTTGTAGTGCAGGTTAAA 1740

QY 1741 AAGCTTAGGCGCCCGAACCAAGCGGAGCTGTTTCTTTTGAAGAAACAGGATAATAC 1800

DB 1741 AAGCTTAGGCGCCCGAACCAAGCGGAGCTGTTTCTTTTGAAGAAACAGGATAATAC 1800

QY 1801 ATGGCGCTTATACGCGCTTCTCCCAACAGACGCGAGGCTTCTTGGTGCATCATCACT 1860



Db 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGGCTTACTTTGGTGTGATCATCACT 1860  
Qy 1861 AGCTTCAAGGCCGGGACAGGAACCAAGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTTCAAGGCCGGGACAGGAACCAAGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Qy 1921 ACACAACTTTCTGCGGACCTGCGTCAATGCGGTGTGTTGACGTGTCTATCATGTGTC 1980  
Db 1921 ACACAACTTTCTGCGGACCTGCGTCAATGCGGTGTGTTGACGTGTCTATCATGTGTC 1980  
Qy 1981 GGCTCAAGAACCTTTGCGGCGCCAAAGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
Db 1981 GGCTCAAGAACCTTTGCGGCGCCAAAGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
Qy 2041 CAGGACCTGCTGCGTGGCAAGCGCCCGCGGGGCGGTCTTGTGACACATGCACTTGC 2100  
Db 2041 CAGGACCTGCTGCGTGGCAAGCGCCCGCGGGGCGGTCTTGTGACACATGCACTTGC 2100  
Qy 2101 GGCAGCTCGGACCTTTACTTTGTTGACGAGCATGCCATGTCAATCCGGTGGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTTGTTGACGAGCATGCCATGTCAATCCGGTGGCGCGCGG 2160  
Qy 2161 GCGACAGAGGGGAGCTACTCTCCCGAGCGCGTCTCTCTTGAAGGCTCTTGC 2220  
Db 2161 GCGACAGAGGGGAGCTACTCTCCCGAGCGCGTCTCTCTTGAAGGCTCTTGC 2220  
Qy 2221 GCGGTTCCACTGCTCTGCGGCGAGCTGTGGGATCTTTGCGGTCGCGTGTGC 2280  
Db 2221 GCGGTTCCACTGCTCTGCGGCGAGCTGTGGGATCTTTGCGGTCGCGTGTGC 2280  
Qy 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Qy 2341 CGGTCCCGGCTTTACGGAACAACCTCGTCCCTCCGCGCGTACCGGAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTACGGAACAACCTCGTCCCTCCGCGCGTACCGGAGACATTCAGGTG 2400  
Qy 2401 GCCATCTACGCGGCTTACTGGTAGGGCAAGACGACTAAGTGGCGGTGGTAGCA 2460  
Db 2401 GCCATCTACGCGGCTTACTGGTAGGGCAAGACGACTAAGTGGCGGTGGTAGCA 2460  
Qy 2461 GCCAAGGTTAAGGTGTCTCTCAACCGGTCCGTCCGCGCACTTAGGTTCGGG 2520  
Db 2461 GCCAAGGTTAAGGTGTCTCTCTCAACCGGTCCGTCCGCGCACTTAGGTTCGGG 2520  
Qy 2521 GCGTATATGTCTAAGGCACATGTPATCGACCTTAACATCAGAACCGGGGTAAAGCACATC 2580  
Db 2521 GCGTATATGTCTAAGGCACATGTPATCGACCTTAACATCAGAACCGGGGTAAAGCACATC 2580  
Qy 2581 ACCACGGGTGCGCCCATCAGTACTCCACTATGGCAAGTTCTTCCGCGAGGTGTTC 2640  
Db 2581 ACCACGGGTGCGCCCATCAGTACTCCACTATGGCAAGTTCTTCCGCGAGGTGTTC 2640  
Qy 2641 TCTGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
Qy 2701 ATCTGGGCGCTATGACATCATATATGTGATGAGTGCCACTCGTGTG 2760  
Db 2701 ATCTGGGCGCTATGACATCATATATGTGATGAGTGCCACTCGTGTG 2760  
Qy 2761 CTGCGCACCGTACGCTCCGGGATCGGTGACCGTGCACATCCAAACATCGAGGAGTG 2820  
Db 2761 CTGCGCACCGTACGCTCCGGGATCGGTGACCGTGCACATCCAAACATCGAGGAGTG 2820  
Qy 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGAAGCCATCCCATCGAGACCATC 2880  
Qy 2881 AAGGGGGGAGGCACTCATTTTCTGCCATTCGAAGAAATGTGATGAGTCTGCGCGG 2940  
Db 2881 AAGGGGGGAGGCACTCATTTTCTGCCATTCGAAGAAATGTGATGAGTCTGCGCGG 2940

Qy 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTCAATGATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTCAATGATCCGTC 3000  
Qy 3001 ATACCAATTAGCGGAGAGCTGATTGTCGTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
Db 3001 ATACCAATTAGCGGAGAGCTGATTGTCGTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
Qy 3061 GCGGATTTTCGACTCAGTGTGATGACTGCAATATGTCCTCACCAGACAGTCCACTTCAGC 3120  
Db 3061 GCGGATTTTCGACTCAGTGTGATGACTGCAATATGTCCTCACCAGACAGTCCACTTCAGC 3120  
Qy 3121 CTGAGCCCGACCTTACCATTGAGACGACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
Db 3121 CTGAGCCCGACCTTACCATTGAGACGACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
Qy 3181 CAGCGGCGAGGAGAGCTGCTAGGGGAGGATGGGCAATTTAGAGTTCTGACTCCAGGA 3240  
Db 3181 CAGCGGCGAGGAGAGCTGCTAGGGGAGGATGGGCAATTTAGAGTTCTGACTCCAGGA 3240  
Qy 3241 GAACGSCCTCGGGCATGTTCAATTCCTCGGTTCTGTGCGAGTCTATGACGCGGCTGT 3300  
Db 3241 GAACGSCCTCGGGCATGTTCAATTCCTCGGTTCTGTGCGAGTCTATGACGCGGCTGT 3300  
Qy 3301 GCTTGTGACGAGCTCACGCGCGCGGAGACCTCAGTTTGTGCGGCTTACCTTAAACACA 3360  
Db 3301 GCTTGTGACGAGCTCACGCGCGCGGAGACCTCAGTTTGTGCGGCTTACCTTAAACACA 3360  
Qy 3361 CAGGGTTCCGCTCTGCGGAGACCTCTGAGTTCTGGGAGAGCGTCTTACAGGCTC 3420  
Db 3361 CAGGGTTCCGCTCTGCGGAGACCTCTGAGTTCTGGGAGAGCGTCTTACAGGCTC 3420  
Qy 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTAAGCAGGAGAGACAACTTCCCTAC 3480  
Db 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTAAGCAGGAGAGACAACTTCCCTAC 3480  
Qy 3481 CTGTTAGCATACAGGCTACGCTGTCGCGAGGGCTCAGGCTCCACTCCCTGTTGGGAC 3540  
Db 3481 CTGTTAGCATACAGGCTACGCTGTCGCGAGGGCTCAGGCTCCACTCCCTGTTGGGAC 3540  
Qy 3541 CAAATGTGAGTGTCTCATACGGCTAAGGCTTACGCTGACGCGGCAACGCGGCTGCTG 3600  
Db 3541 CAAATGTGAGTGTCTCATACGGCTAAGGCTTACGCTGACGCGGCAACGCGGCTGCTG 3600  
Qy 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACACCCCAATAACATAATC 3660  
Db 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACACCCCAATAACATAATC 3660  
Qy 3661 ATGCGCATGATGTCGCTGACCTGGAGGTCGTCAAGACACCTGGGTGCTGTAGGCGGA 3720  
Db 3661 ATGCGCATGATGTCGCTGACCTGGAGTCTGTCAGAGCACCTGGGTGCTGTAGGCGGA 3720  
Qy 3721 GTCTAGCAGCTGTGGCGGCTATTCGCTGACACAGCAGCGGTGCTATTGTGGGAGG 3780  
Db 3721 GTCTAGCAGCTGTGGCGGCTATTCGCTGACACAGCAGCGGTGCTATTGTGGGAGG 3780  
Qy 3781 ATCATCTTGTCCGAAAGCGGCTCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCTCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Qy 3841 GATGAGATGGAAGAGTGGCTCTCACCTCCCTTACATCGAAACAGGGAATGAGCTCGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGCTCTCACCTCCCTTACATCGAAACAGGGAATGAGCTCGCC 3900  
Qy 3901 GAACAAATCAACAGAGGCAATCGGTTCTGCTCAACACAGCCCAAGAGCGGAGGCT 3960  
Db 3901 GAACAAATCAACAGAGGCAATCGGTTCTGCTCAACACAGCCCAAGAGCGGAGGCT 3960  
Qy 3961 GCTCTCCGTTGGTGAATCCAGTGGCGGACCTCGAAGGCTTCTGGCGGAAGCATATG 4020  
Db 3961 GCTCTCCGTTGGTGAATCCAGTGGCGGACCTCGAAGGCTTCTGGCGGAAGCATATG 4020



QY 4021 TGGAAATTCATCAGCGGGATACAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGGATACAATATTTAGAGGCTTGTCCACTCTGCTGGCAACCC 4080  
QY 4081 GGGATAGCATCACTGATGCGATTCACAGCCTCTATCACCAGCGCGTCAACCAACAT 4140  
Db 4081 GGGATAGCATCACTGATGCGATTCACAGCCTCTATCACCAGCGCGTCAACCAACAT 4140  
QY 4141 ACCCTCCTGTTTAAATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCCTGTTTAAATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTCGTAGGCGCGCATGCTGGAGCGGCTGTGGCAGCATAGGCCCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCGCATGCTGGAGCGGCTGTGGCAGCATAGGCCCTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCGTGCC 4320  
Db 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCGTGCC 4320  
QY 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCCTCCTGCT 4380  
Db 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTCCTCCTGCT 4380  
QY 4381 ATCCTCTCCCTGGCGCCTTAGTCGTCGGGTCTGTGGCAGCGATACCTGCTCGGCAC 4440  
Db 4381 ATCCTCTCCCTGGCGCCTTAGTCGTCGGGTCTGTGGCAGCGATACCTGCTCGGCAC 4440  
QY 4441 GTGGGCCAGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCCTTCCGTTTCGGG 4500  
Db 4441 GTGGGCCAGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCCTTCCGTTTCGGG 4500  
QY 4501 GGTAAACACAGCTCTCCCGCAGCACTATGTGCTGAGCGGAGCTGTGACGAGCTGTCACT 4560  
Db 4501 GGTAAACACAGCTCTCCCGCAGCACTATGTGCTGAGCGGAGCTGTGACGAGCTGTCACT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACATCACTCAGCTGTGTAAGAGGCTTCAACAGTGAATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACATCACTCAGCTGTGTAAGAGGCTTCAACAGTGAATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTGGATTTGGATATGC 4680  
Db 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTGGATTTGGATATGC 4680  
QY 4681 ACGGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGCTCCTGCGCGATTCGCCGGA 4740  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGCTCCTGCGCGATTCGCCGGA 4740  
QY 4741 GTCCCTTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGCGGACGATCATG 4800  
Db 4741 GTCCCTTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGCGGACGATCATG 4800  
QY 4801 CAACACACCTGCCATGTGGAGCACAGATCACCGGACATGTGAAAAACGTTCCATGAGG 4860  
Db 4801 CAACACACCTGCCATGTGGAGCACAGATCACCGGACATGTGAAAAACGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACCTGCGATGCAATTCGCCATTAACGCTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAAACCTGCGATGCAATTCGCCATTAACGCTAC 4920  
QY 4921 ACCACGGGCCCCGACGCCCTCCCGCGGCCAAATTTATTCTAGGGCGCTGTGGCGGGTG 4980  
Db 4921 ACCACGGGCCCCGACGCCCTCCCGCGGCCAAATTTATTCTAGGGCGCTGTGGCGGGTG 4980  
QY 4981 GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTCCACTACGTGAGCGGGCATG 5040  
Db 4981 GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTCCACTACGTGAGCGGGCATG 5040  
QY 5041 ACCACTGACAACTGAAGTCCCGTGTGACGTTCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
Db 5041 ACCACTGACAACTGAAGTCCCGTGTGACGTTCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
QY 5101 GATGGGCTGGGTTGACAGGTACGCTCCAGCGTGCMAACCCCTCTACGGGAGGAGTTC 5160

Db 5101 GATGGGCTGGGTTGACAGGTACGCTCCAGCGTGCMAACCCCTCTACGGGAGGAGTTC 5160  
QY 5161 ACATTCCTGGTCCGGCTCAATCAATACCTGTTGGTTCACAGTCCCATGCGAGCCCGAA 5220  
Db 5161 ACATTCCTGGTCCGGCTCAATCAATACCTGTTGGTTCACAGTCCCATGCGAGCCCGAA 5220  
QY 5221 CCGGAGTAGCATGCTCACTTCCATGTCTACCGACCCCTCCACATTAAGCGGAGACG 5280  
Db 5221 CCGGAGTAGCATGCTCACTTCCATGTCTACCGACCCCTCCACATTAAGCGGAGACG 5280  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTGAGTCCCGGAGCGCTGAC 5340  
Db 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTGAGTCCCGGAGCGCTGAC 5340  
QY 5341 CTGTCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGCTCCCGGAGCGCTGAC 5400  
Db 5341 CTGTCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGCTCCCGGAGCGCTGAC 5400  
QY 5401 CTCATCGAGGCAACCTCTCTGCGGCGAGGATGGCGGGAAATCACCCGCGTGGAG 5460  
Db 5401 CTCATCGAGGCAACCTCTCTGCGGCGAGGATGGCGGGAAATCACCCGCGTGGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTCGAGCGCTTCCAGCGGAGGATGAG 5520  
Db 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTCGAGCGCTTCCAGCGGAGGATGAG 5520  
QY 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCACAGAAATTCCTCGAGCGATG 5580  
Db 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCACAGAAATTCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGCAACCCCGGATTAACAACCTTCTAGAGTCTTGGAGGACCCGAC 5640  
Db 5581 CCCATATGGGCAACCCCGGATTAACAACCTTCTAGAGTCTTGGAGGACCCGAC 5640  
QY 5641 TACGTCCTCCAGTGTACAGGCTGCTTTCGCGCTGCTGCAAGGCGCTTCCGATACCA 5700  
Db 5641 TACGTCCTCCAGTGTACAGGCTGCTTTCGCGCTGCTGCAAGGCGCTTCCGATACCA 5700  
QY 5701 CCTCCAGGAGGAGGACGTTTGTCTGTGCAATCTACCGTGTCTTCTGCTTGGCG 5760  
Db 5701 CCTCCAGGAGGAGGACGTTTGTCTGTGCAATCTACCGTGTCTTCTGCTTGGCG 5760  
QY 5761 GAGTCTGCCACAAAGACCTTTCGCGAGCTCCGAATCTGCGCGCTGCGACGCGCACGCA 5820  
Db 5761 GAGTCTGCCACAAAGACCTTTCGCGAGCTCCGAATCTGCGCGCTGCGACGCGCACGCA 5820  
QY 5821 ACGGCTCTCTGACGAGCGCTTCCGAGCGGAGTCCGAGCGTGGATCGGATCGTAC 5880  
Db 5821 ACGGCTCTCTGACGAGCGCTTCCGAGCGGAGTCCGAGCGTGGATCGGATCGTAC 5880  
QY 5881 TCCTCCATGCCCCCTTGAAGGGGAGCGCGGATTCGCCATCTCAGCGAGGGTCTTGG 5940  
Db 5881 TCCTCCATGCCCCCTTGAAGGGGAGCGCGGATTCGCCATCTCAGCGAGGGTCTTGG 5940  
QY 5941 TCTACCGTAAAGCAGGAGGCTAGTGGAGGCTGCTGTGCTGCTGATGTCTTACATG 6000  
Db 5941 TCTACCGTAAAGCAGGAGGCTAGTGGAGGCTGCTGTGCTGCTGATGTCTTACATG 6000  
QY 6001 ACAGGCGCTGATCAGCGCATGCGTGGGAGGAAACAAAGTGGCCCATCAATGACATG 6060  
Db 6001 ACAGGCGCTGATCAGCGCATGCGTGGGAGGAAACAAAGTGGCCCATCAATGACATG 6060  
QY 6061 AGCAACTCTTGTCTCCGTACCACTTGTGTATGCTACACATCTCGCAGCGCAAGC 6120  
Db 6061 AGCAACTCTTGTCTCCGTACCACTTGTGTATGCTACACATCTCGCAGCGCAAGC 6120  
QY 6121 CTGCGGAGAGGATCACTTTTGAAGACTGCGGTCCTTGGACGACCACTTACCGGAG 6180  
Db 6121 CTGCGGAGAGGATCACTTTTGAAGACTGCGGTCCTTGGACGACCACTTACCGGAG 6180  
QY 6181 GTGCTCAAGGATGAAGCGGAGGCTCCAAGTTAAGGCTAAATCTTATCCGTSAG 6240

Db 6181 GTGCTCAGGAGATGAAGCGGAGCGCTCCACAGTTAAGGCTAATACTTCTATCCGTGGAG 6240  
Qy 6241 GAAGCCTGTAAGCTGACGCCGCCCAANTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6300  
Db 6241 GAAGCCTGTAAGCTGACGCCGCCCAANTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6300  
Qy 6301 GACGTCCGGAACCTATCCAGCAGCGCGTTAACCAATCCGCTCGGTGGGAAGACTTG 6360  
Db 6301 GACGTCCGGAACCTATCCAGCAGCGCGTTAACCAATCCGCTCGGTGGGAAGACTTG 6360  
Qy 6361 CTGGAAGACACTGAGACACCAATTCACACCAATCATGCGCAAAAATGAGTTTTCTGC 6420  
Db 6361 CTGGAAGACACTGAGACACCAATTCACACCAATCATGCGCAAAAATGAGTTTTCTGC 6420  
Qy 6421 GTTCAACACGAGAAGCGGGCGCGAAGCAGCTCGCTTTATCGTATTCACGATTTGGG 6480  
Db 6421 GTTCAACACGAGAAGCGGGCGCGAAGCAGCTCGCTTTATCGTATTCACGATTTGGG 6480  
Qy 6481 GTTGTGTGTGCGAGAAATGGCCCTTTACGATGTGCTCCACCCCTCCCTCAGGCCGTG 6540  
Db 6481 GTTGTGTGTGCGAGAAATGGCCCTTTACGATGTGCTCCACCCCTCCCTCAGGCCGTG 6540  
Qy 6541 ATGGGCTCTTACGATTTCAATCTCTCTGGACAGCGGTCGAGTTCTCTGTGAAT 6600  
Db 6541 ATGGGCTCTTACGATTTCAATCTCTCTGGACAGCGGTCGAGTTCTCTGTGAAT 6600  
Qy 6601 GCCTGAAACGAGAAATGCCCTATGGGCTTCGCAATGACACCCGCTGTTGTGACTCA 6660  
Db 6601 GCCTGAAACGAGAAATGCCCTATGGGCTTCGCAATGACACCCGCTGTTGTGACTCA 6660  
Qy 6661 ACGTCACTGAGATGACATCCGCTGTTGAGGATCAATCTACCAATGTTGACTTGGCC 6720  
Db 6661 ACGTCACTGAGATGACATCCGCTGTTGAGGATCAATCTACCAATGTTGACTTGGCC 6720  
Qy 6721 CCGGAGACGACAGAGCCATTAAGCTCGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG 6780  
Db 6721 CCGGAGACGACAGAGCCATTAAGCTCGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG 6780  
Qy 6781 ACTAATTTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Db 6781 ACTAATTTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Qy 6841 ACCAGTGGGTAATACCTTCATGTTACTTCAAGCGCGCTCGGCTGCTGAGGCTGCG 6900  
Db 6841 ACCAGTGGGTAATACCTTCATGTTACTTCAAGCGCGCTCGGCTGCTGAGGCTGCG 6900  
Qy 6901 AAGCTCCAGACTGACAGATGCTGATGCGGAGACGACCTTGTCTGTTACTGTGAAGC 6960  
Db 6901 AAGCTCCAGACTGACAGATGCTGATGCGGAGACGACCTTGTCTGTTACTGTGAAGC 6960  
Qy 6961 GCGGGACCCAAAGAGGAGGAGCGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAAAGAGGAGGAGCGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Qy 7021 TCTGCCCTTGGGAGCGGCGCCCAACACGATACGACTTGGAGTTGATAACATATGC 7080  
Db 7021 TCTGCCCTTGGGAGCGGCGCCCAACACGATACGACTTGGAGTTGATAACATATGC 7080  
Qy 7081 TCCTCCAAATGTGTGAGTCCGCGAGATGCTGCTGCAAAAGGGTGTACTATCTCACCGT 7140  
Db 7081 TCCTCCAAATGTGTGAGTCCGCGAGATGCTGCTGCAAAAGGGTGTACTATCTCACCGT 7140  
Qy 7141 GACCCACACCCCTTGGCGGCTGCGTGGAGACAGTACACACACTCCAGTCAAT 7200  
Db 7141 GACCCACACCCCTTGGCGGCTGCGTGGAGACAGTACACACACTCCAGTCAAT 7200  
Qy 7201 TCCTGGCTAGGACATCATGATGCGCCCTTGTGGCAAGGATGATCTGATG 7260  
Db 7201 TCCTGGCTAGGACATCATGATGCGCCCTTGTGGCAAGGATGATCTGATG 7260  
Qy 7261 ACTCATTTCTTCTCCTCCTTCTAGCTCAGGACAACTTTGAAAGCCCTAGATTGTCAG 7320  
Db 7261 ACTCATTTCTTCTCCTCCTTCTAGCTCAGGACAACTTTGAAAGCCCTAGATTGTCAG 7320

Qy 7321 ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACTCTCAGATCATTCACGACTC 7380  
Db 7321 ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACTCTCAGATCATTCACGACTC 7380  
Qy 7381 CATGCCCTTAGCGCATTTTCTACTCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
Db 7381 CATGCCCTTAGCGCATTTTCTACTCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
Qy 7441 TCATGCCCTCAGAAACTTTGGGTACCGCCCTTGCGAGTCTGAGACATCGGGCCAGAACT 7500  
Db 7441 TCATGCCCTCAGAAACTTTGGGTACCGCCCTTGCGAGTCTGAGACATCGGGCCAGAACT 7500  
Qy 7501 GTCCGCGCTTAGCTACTGTCTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGCGCTTAGCTACTGTCTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Qy 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAACTCCGCTCGCTCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAACTCCGCTCGCTCCAGTTGGAT 7620  
Qy 7621 TTATCCAGCTGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7680  
Db 7621 TTATCCAGCTGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7680  
Qy 7681 GCCCGACCCGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7740  
Db 7681 GCCCGACCCGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7740  
Qy 7741 CTACTCCCCAAACGATGAAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCCAAACGATGAAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Qy 7801 TTTCCCTTT 7860  
Db 7801 TTTCCCTTT 7860  
Qy 7861 TTTTCTCTCTTT 7920  
Db 7861 TTTTCTCTCTTT 7920  
Qy 7921 TAGCTGTGAAGTCCGCTGAGCCGCTGAGTCTGAGAGAGTGTGATAGTGGCTCTCTGC 7980  
Db 7921 TAGCTGTGAAGTCCGCTGAGCCGCTGAGTCTGAGAGAGTGTGATAGTGGCTCTCTGC 7980  
Qy 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

## RESULT 8

US-10-434-842-6

; Sequence 6, Application US/10434842

; Publication No. US2004005549A1

; GENERAL INFORMATION:

; APPLICANT: Bichko, Vadim

; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE

; FILE REFERENCE: 0342/1H395US3

; CURRENT APPLICATION NUMBER: US/10/434,842

; PRIOR FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: US 10/233,307

; PRIOR FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: US 10/005,469

; PRIOR FILING DATE: 2001-11-07

; PRIOR APPLICATION NUMBER: US 60/245,866

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 7992

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

OTHER INFORMATION: HCV24 subgenomic HCV replicon  
US-10-434-842-6

Query Match	100.0%	Score 7985.8	DB 16	Length 7992
Best Local Similarity	100.0%	Pred. No. 0		
Matches 7987	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 1	GCACGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTCGTGGAGAACTACTG	60		
DB 1	GCACGCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTCGTGGAGAACTACTG	60		
QY 61	TCTTCACGAGAAACGCTAGCCATGCGTGTAGTATGAGTGTCTGAGCCTCCAGGAC	120		
DB 61	TCTTCACGAGAAACGCTAGCCATGCGTGTAGTATGAGTGTCTGAGCCTCCAGGAC	120		
QY 121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGAAATGGCCAG	180		
DB 121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGAAATGGCCAG	180		
QY 181	GACGACCGGGTCTCTTTCTTGATCAACCCGCTCAATGCTGAGATTTGGCGTCCCCC	240		
DB 181	GACGACCGGGTCTCTTTCTTGATCAACCCGCTCAATGCTGAGATTTGGCGTCCCCC	240		
QY 241	CGGAGACTGCTAGCGAGTAGTGTGGTTCGGAAGGCTTGTGTACTGCTGATAGG	300		
DB 241	CGGAGACTGCTAGCGAGTAGTGTGGTTCGGAAGGCTTGTGTACTGCTGATAGG	300		
QY 301	GTGCTTGGAGTGTCCCGGGAGGTCTGTAGACGTGCACATGAGCAGCAATCCTAAAC	360		
DB 301	GTGCTTGGAGTGTCCCGGGAGGTCTGTAGACGTGCACATGAGCAGCAATCCTAAAC	360		
QY 361	CTCAAGAAAACCAAGGCGCGCCATGATTGAACAGATGGATTCACACAGCTTCTC	420		
DB 361	CTCAAGAAAACCAAGGCGCGCCATGATTGAACAGATGGATTCACACAGCTTCTC	420		
QY 421	CGGCGGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAGCAATCGGCTGCT	480		
DB 421	CGGCGGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAGCAATCGGCTGCT	480		
QY 481	CTGATGCGCGGCTTCCGGCTGTGAGGCGAGGCGCGCGTCTTTTGTCAAGACCG	540		
DB 481	CTGATGCGCGGCTTCCGGCTGTGAGGCGAGGCGCGCGTCTTTTGTCAAGACCG	540		
QY 541	ACCTGTCCGGTCCCTCAATGAACCTGCAGACAGGCGAGCGGCTATCGTGGCTGSCCA	600		
DB 541	ACCTGTCCGGTCCCTCAATGAACCTGCAGACAGGCGAGCGGCTATCGTGGCTGSCCA	600		
QY 601	CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC	660		
DB 601	CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC	660		
QY 661	TGCTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGA	720		
DB 661	TGCTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCTCATCTCACTTGTCTCTGCCGAGA	720		
QY 721	AAGTATCCATCATGGTGTGATGCAATGCGGCTGTGATAGCTGTGATCCGCTACCTGCC	780		
DB 721	AAGTATCCATCATGGTGTGATGCAATGCGGCTGTGATAGCTGTGATCCGCTACCTGCC	780		
QY 781	CATTTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCACTGCTGGATGGAAGCGGTC	840		
DB 781	CATTTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCACTGCTGGATGGAAGCGGTC	840		
QY 841	TTGTTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG	900		
DB 841	TTGTTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG	900		
QY 901	CCAGGCTCAAGGCGCGCATCCCGACGCGAGGATCTCGTGTGACCCATGGCGATGCTCT	960		
DB 901	CCAGGCTCAAGGCGCGCATCCCGACGCGAGGATCTCGTGTGACCCATGGCGATGCTCT	960		
QY 961	GCTTGGCGAATATCATGGTGGAAATGGCGCGCTTTTCTGGATTCATCGACTGTGGCGGC	1020		
DB 961	GCTTGGCGAATATCATGGTGGAAATGGCGCGCTTTTCTGGATTCATCGACTGTGGCGGC	1020		

Db	961	GCTTGGCGAATATCATGGTGGAAATGGCGCGCTTTTCTGGATTCATCGACTGTGGCGGC	1020
QY	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTTGCTGAAGAGC	1080
Db	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTTGCTGAAGAGC	1080
QY	1081	TTGGCGGGAATGGGCTGACCGTTCCTCGTGTCTTACGGTATCGCGCTCCCGATTGCG	1140
Db	1081	TTGGCGGGAATGGGCTGACCGTTCCTCGTGTCTTACGGTATCGCGCTCCCGATTGCG	1140
QY	1141	AGCGCATCGCTTCTATCGCCTTCTGACGAGTCTTCTGAGTCTTAAACAGACCAACAG	1200
Db	1141	AGCGCATCGCTTCTATCGCCTTCTGACGAGTCTTCTGAGTCTTAAACAGACCAACAG	1200
QY	1201	GTTTCCCTCTAGCGGATCAATTCGCCCTCTCCCTCCCGCCCCCTTAAAGTACTGCG	1260
Db	1201	GTTTCCCTCTAGCGGATCAATTCGCCCTCTCCCTCCCGCCCCCTTAAAGTACTGCG	1260
QY	1261	CGAAGCGCTTGGATTAAGCGCGGTGCGTGTCTATATGTTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGCTTGGATTAAGCGCGGTGCGTGTCTATATGTTATTTTCCACCATATTG	1320
QY	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGSCCTCTCTTCTGACGAGCATTCCT	1380
Db	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGSCCTCTCTTCTGACGAGCATTCCT	1380
QY	1381	AGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCTGTAAGAGCA	1440
Db	1381	AGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCTGTAAGAGCA	1440
QY	1441	GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTAGCGACCTTTTGACGAGCGG	1500
Db	1441	GTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTAGCGACCTTTTGACGAGCGG	1500
QY	1501	AACCCCGACCTGCGGACAGTGTCTCGGCGCAAAAGCAAGCCTGTATAGATACACCT	1560
Db	1501	AACCCCGACCTGCGGACAGTGTCTCGGCGCAAAAGCAAGCCTGTATAGATACACCT	1560
QY	1561	GCAAGGCGGCAACACCCAGTGTGAGTGTGAGTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGGCGGCAACACCCAGTGTGAGTGTGGAAGAGTCAAA	1620
QY	1621	TGGCTCTCTCAAGGATTAACAAGGCTGAGGATGCCAAGAGTCAACCAATTTG	1680
Db	1621	TGGCTCTCTCAAGGATTAACAAGGCTGAGGATGCCAAGAGTCAACCAATTTG	1680
QY	1681	ATGGATCTCATCTGGGCGCTCGGTGCACATGTTTACATGTGTGTTAGTCGAGGTTAAA	1740
Db	1681	ATGGATCTCATCTGGGCGCTCGGTGCACATGTTTACATGTGTGTTAGTCGAGGTTAAA	1740
QY	1741	AACGTCTAGGCCCCCGAACACGCGGACGTTGTTTCTTTGAAAACACGATAATACC	1800
Db	1741	AACGTCTAGGCCCCCGAACACGCGGACGTTGTTTCTTTGAAAACACGATAATACC	1800
QY	1801	ATGGCGCTATTAGGCTACTTCCACAGACGCGAGGCTCTCTGCTGCATCATCACT	1860
Db	1801	ATGGCGCTATTAGGCTACTTCCACAGACGCGAGGCTCTCTGCTGCATCATCACT	1860
QY	1861	AGGCTCACAGGCGGACAGGACACAGGCTCGAGGGAGGTCCCAAGTGTCTCCACCGCA	1920
Db	1861	AGGCTCACAGGCGGACAGGACACAGGCTCGAGGGAGGTCCCAAGTGTCTCCACCGCA	1920
QY	1921	ACAAATCTTTCTGCGGACCTGCTCAATGGCGTGTGTTGGACTGTCTATCATGTTGCC	1980
Db	1921	ACAAATCTTTCTGCGGACCTGCTCAATGGCGTGTGTTGGACTGTCTATCATGTTGCC	1980
QY	1981	GGCTCAAGACCTTTCGCGGCGCAAGGCGCCAAATCACCAATGTACCAATGTGGAC	2040
Db	1981	GGCTCAAGACCTTTCGCGGCGCAAGGCGCCAAATCACCAATGTACCAATGTGGAC	2040
QY	2041	CAGGACCTCGTGGTGGCAAGGCGCGCGCGGCGGCTTCTTGACACATGACCTGC	2100
Db	2041	CAGGACCTCGTGGTGGCAAGGCGCGCGCGGCGGCTTCTTGACACATGACCTGC	2100

QY 2101 GCGAGCTCGGACCTTTTACTGGTACAGAGGAGATGCGATGTCATTTCCGGTGGCCGCGG 2160  
 Db 2101 GCGAGCTCGGACCTTTTACTGGTACAGAGGAGATGCGATGTCATTTCCGGTGGCCGCGG 2160  
 QY 2161 GCGGACAGAGGGAGGACTACTCTCCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
 Db 2161 GCGGACAGAGGGAGGACTACTCTCCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
 QY 2221 GCGGCTCCACTCTCTGCGCCCTCGGGGACAGCTGTGGGCATCTTTCCGGGCTCCGCTGTC 2280  
 Db 2221 GCGGCTCCACTCTCTGCGCCCTCGGGGACAGCTGTGGGCATCTTTCCGGGCTCCGCTGTC 2280  
 QY 2281 ACCCGAGGGGTTGCGAAGGGCGGTGAGCTTTGTGTAACCGTCTGATGTAAGAAACCACTATG 2340  
 Db 2281 ACCCGAGGGGTTGCGAAGGGCGGTGAGCTTTGTGTAACCGTCTGATGTAAGAAACCACTATG 2340  
 QY 2341 CGGTCCCGGCTCTTCAAGGACAACTCGTCCCTCGGGCGGTACCGGAGACATTTCCAGGTG 2400  
 Db 2341 CGGTCCCGGCTCTTCAAGGACAACTCGTCCCTCGGGCGGTACCGGAGACATTTCCAGGTG 2400  
 QY 2401 GCGCATCTACAGCCCTTACTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
 Db 2401 GCGCATCTACAGCCCTTACTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
 QY 2461 GCGCAAGGGTATAAGGTGCTTGTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2520  
 Db 2461 GCGCAAGGGTATAAGGTGCTTGTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2520  
 QY 2521 GGTATATCTTAAGGACATGTAATCGACCTTAACATCAGAACCGGGTAAAGACCATC 2580  
 Db 2521 GGTATATCTTAAGGACATGTAATCGACCTTAACATCAGAACCGGGTAAAGACCATC 2580  
 QY 2581 ACCAGGGTGCCTCCATCAGTACTCCACTATGGAAGTTTCTCCGACGGTGTGTC 2640  
 Db 2581 ACCAGGGTGCCTCCATCAGTACTCCACTATGGAAGTTTCTCCGACGGTGTGTC 2640  
 QY 2641 TCTGGGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCAT 2700  
 Db 2641 TCTGGGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCAT 2700  
 QY 2701 ATCTGGGATCGGACAGTCTGGACCAAGCGGAGAGCGCTGAGCGGACCTCTGCTG 2760  
 Db 2701 ATCTGGGATCGGACAGTCTGGACCAAGCGGAGAGCGCTGAGCGGACCTCTGCTG 2760  
 QY 2761 CTGCGCACCGCTACGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG 2820  
 Db 2761 CTGCGCACCGCTACGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG 2820  
 QY 2821 GCTCTGTGAGGACTCGGAGAAATCCCTTTTATGCGAAGCAATCCCATCGAGACCATC 2880  
 Db 2821 GCTCTGTGAGGACTCGGAGAAATCCCTTTTATGCGAAGCAATCCCATCGAGACCATC 2880  
 QY 2881 AAGGGGGAGGACACTCATTTTCTGCAATCCAAAGAAATGATGATGAGCTCGCGCG 2940  
 Db 2881 AAGGGGGAGGACACTCATTTTCTGCAATCCAAAGAAATGATGATGAGCTCGCGCG 2940  
 QY 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATATCGCTC 3000  
 Db 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATATCGCTC 3000  
 QY 3001 ATACCAACTAGCGGAGCGTCTATGCTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
 Db 3001 ATACCAACTAGCGGAGCGTCTATGCTAGCAACGAGCGCTCTAATGACGGGCTTTACC 3060  
 QY 3061 GCGGATTTGACTCAGTATGCTGCAATACATGTCACCCAGACAGTCTGACTCAGC 3120  
 Db 3061 GCGGATTTGACTCAGTATGCTGCAATACATGTCACCCAGACAGTCTGACTCAGC 3120  
 QY 3121 CTGACCGGACCTTCACTATGAGACGACGACCGGTGCGCAAGACGGGTGTACGCTCG 3180  
 Db 3121 CTGACCGGACCTTCACTATGAGACGACGACCGGTGCGCAAGACGGGTGTACGCTCG 3180

QY 3181 CAGCGGCGAGCGAGGACTGGTAGGCGCAGGATGGGCAATTTACAGTTTGTGTGACTCAGGA 3240  
 Db 3181 CAGCGGCGAGCGAGGACTGGTAGGCGCAGGATGGGCAATTTACAGTTTGTGTGACTCAGGA 3240  
 QY 3241 GAAACGGCCCTCGGGCATGTTCCGTTCTGCGGTTCTGCGGAGTGTATGACCGGGCTGT 3300  
 Db 3241 GAAACGGCCCTCGGGCATGTTCCGTTCTGCGGTTCTGCGGAGTGTATGACCGGGCTGT 3300  
 QY 3301 GCTTGGTACGAGCTCAGCGCCGCGAGACCTCAGTTAGTTTCCGGCTTTACCTAAACACA 3360  
 Db 3301 GCTTGGTACGAGCTCAGCGCCGCGAGACCTCAGTTAGTTTCCGGCTTTACCTAAACACA 3360  
 QY 3361 CCAGGGTTGCCGCTGTCAGGACCATCTGGAGTTCTGGGAGAGCTCTTTACAGGCTTC 3420  
 Db 3361 CCAGGGTTGCCGCTGTCAGGACCATCTGGAGTTCTGGGAGAGCTCTTTACAGGCTTC 3420  
 QY 3421 ACCCAATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGAGCACTTCCCTTAC 3480  
 Db 3421 ACCCAATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGAGCACTTCCCTTAC 3480  
 QY 3481 CTGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
 Db 3481 CTGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
 QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGCGCAACGCGCTGCTG 3600  
 Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGCGCAACGCGCTGCTG 3600  
 QY 3601 TATAGGCTGGAGCGGTTCAAAAAGAGTTACTACACACACCCCATACCAANTACATC 3660  
 Db 3601 TATAGGCTGGAGCGGTTCAAAAAGAGTTACTACACACACCCCATACCAANTACATC 3660  
 QY 3661 ATGGCATGATGTGCGCTCAGCTGGAGTGTCTCACGAGCACCTGGGTGTGTAGGCGGA 3720  
 Db 3661 ATGGCATGATGTGCGCTCAGCTGGAGTGTCTCACGAGCACCTGGGTGTGTAGGCGGA 3720  
 QY 3721 GTCTAGGAGCTGTGCGCGCTATTCCTGCAACAGGAGCGTGTGTATGTTGGGCGG 3780  
 Db 3721 GTCTAGGAGCTGTGCGCGCTATTCCTGCAACAGGAGCGTGTGTATGTTGGGCGG 3780  
 QY 3781 ATCATCTGTTCGGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
 Db 3781 ATCATCTGTTCGGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
 QY 3841 GATGAGATGGAAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCACTCGCC 3900  
 Db 3841 GATGAGATGGAAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCACTCGCC 3900  
 QY 3901 GAAACAAATTCAAAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGT 3960  
 Db 3901 GAAACAAATTCAAAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGT 3960  
 QY 3961 GCTGCTCCCGTGTGGAATTCAGAGTGGCGGACCTCGAGGCTTCTGGGCGAGCATATG 4020  
 Db 3961 GCTGCTCCCGTGTGGAATTCAGAGTGGCGGACCTCGAGGCTTCTGGGCGAGCATATG 4020  
 QY 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC 4080  
 Db 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC 4080  
 QY 4081 GCGATAGCATCTGATGGAATTCAGAGCTCTTATACAGGCGGCTTCAACCCACAT 4140  
 Db 4081 GCGATAGCATCTGATGGAATTCAGAGCTCTTATACAGGCGGCTTCAACCCACAT 4140  
 QY 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCACTTGTCTCTCCAGGCT 4200  
 Db 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCACTTGTCTCTCCAGGCT 4200  
 QY 4201 GCTTCTCTTTCGTAGGCGCGGCTCGCTGGAGCGGCTGTGTCAGCATAGGCTTGGG 4260  
 Db 4201 GCTTCTCTTTCGTAGGCGCGGCTCGCTGGAGCGGCTGTGTCAGCATAGGCTTGGG 4260  
 QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTTGGAGCAGGGGTGGCAGGCGCTCGTGCC 4320

Db 4261 AAGGTGCTGTGGATATTTGGCAGGTTATGAGCAGGGGTGCAGGCGCGTCTGTGGCC 4320  
Qy |||||  
Db 4321 TTTAAGGTTCATCAGCGGCGAGTGCCTCCACCGAGGACCTGGTTAA CTTACTCCCTGCT 4380  
Qy |||||  
Db 4321 TTTAAGGTTCATCAGCGGCGAGTGCCTCCACCGAGGACCTGGTTAA CTTACTCCCTGCT 4380  
Qy |||||  
Db 4381 APTCTCTCCCTTGGGCGCCCTAGTCGTGCGGGTCTGTGTGGCAGCGGATATGCGTGGGCAC 4440  
Qy |||||  
Db 4381 ATCTCTCTCCCTTGGGCGCCCTAGTCGTGCGGGTCTGTGTGGCAGCGGATATGCGTGGGCAC 4440  
Qy |||||  
Db 4441 GTGGGCCACAGGGGAGGGGCTGTGTGAGTGGATGAACCCGCTGATAGCGTTCCCTTCCGGG 4500  
Qy |||||  
Db 4441 GTGGGCCACAGGGGAGGGGCTGTGTGAGTGGATGAACCCGCTGATAGCGTTCCCTTCCGGG 4500  
Qy |||||  
Db 4501 GGTAAACACAGTCTCCCAACGCACTATGTGCTGAGAGGAGCGTGCAGCGTGTCACT 4560  
Qy |||||  
Db 4501 GGTAAACACAGTCTCCCAACGCACTATGTGCTGAGAGGAGCGTGCAGCGTGTCACT 4560  
Qy |||||  
Db 4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGGCTTACCAGTGGATCAAC 4620  
Qy |||||  
Db 4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGGCTTACCAGTGGATCAAC 4620  
Qy |||||  
Db 4621 GAGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAGAGATGTTGGGATTTGGATATGC 4680  
Qy |||||  
Db 4621 GAGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAGAGATGTTGGGATTTGGATATGC 4680  
Qy |||||  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTCTGCGCGGATTCGCGGA 4740  
Qy |||||  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTCTGCGCGGATTCGCGGA 4740  
Qy |||||  
Db 4741 GTCCCTCTCTCATGTCACAGTGGGTACAAAGAGTCTGCGGGGGAGCGGATCATG 4800  
Qy |||||  
Db 4741 GTCCCTCTCTCATGTCACAGTGGGTACAAAGAGTCTGCGGGGGAGCGGATCATG 4800  
Qy |||||  
Db 4801 CAACACACCTGCTGAGAGCAAGATCAACCGGACATGTGAAAAAGCTTCCATGAGG 4860  
Qy |||||  
Db 4801 CAACACACCTGCTGAGAGCAAGATCAACCGGACATGTGAAAAAGCTTCCATGAGG 4860  
Qy |||||  
Db 4861 ATCTGCGGGCTAGGACTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACCGGTAC 4920  
Qy |||||  
Db 4861 ATCTGCGGGCTAGGACTGTAGTAAACAGTGGGATGGAACATTTCCCATTAACCGGTAC 4920  
Qy |||||  
Db 4921 ACCACGGGCCCCCTGACGCCCTCCCGCGGCGCAAAATTTATCTAGGCGGCTGTGCGGGTG 4980  
Qy |||||  
Db 4921 ACCACGGGCCCCCTGACGCCCTCCCGCGGCGCAAAATTTATCTAGGCGGCTGTGCGGGTG 4980  
Qy |||||  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Qy |||||  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Qy |||||  
Db 5041 ACCACTGACAACTGAAGTGCCTGTACAGTTCCCGGCCCGGAAATTTCTTCAAGAGTG 5100  
Qy |||||  
Db 5041 ACCACTGACAACTGAAGTGCCTGTACAGTTCCCGGCCCGGAAATTTCTTCAAGAGTG 5100  
Qy |||||  
Db 5101 GATGGGGTGGGTTGCACAGGTACGCTCCAGCGTGAACCCCTCTACCGGAGGAGGTC 5160  
Qy |||||  
Db 5101 GATGGGGTGGGTTGCACAGGTACGCTCCAGCGTGAACCCCTCTACCGGAGGAGGTC 5160  
Qy |||||  
Db 5161 ACATTCCTGTGGGTCAATCAATACCTGTTGGGTACAGCTCCCATCCGAGCCCGGAA 5220  
Qy |||||  
Db 5161 ACATTCCTGTGGGTCAATCAATACCTGTTGGGTACAGCTCCCATCCGAGCCCGGAA 5220  
Qy |||||  
Db 5221 CCGGAGTACAGTGTCACTTCCAGTCCAGGACCCCTCCACATTAAGCGGAGAGC 5280  
Qy |||||  
Db 5221 CCGGAGTACAGTGTCACTTCCAGTCCAGGACCCCTCCACATTAAGCGGAGAGC 5280  
Qy |||||  
Db 5281 GCTAAGCGTAGGCTGCGCAGGGGATCTCCCGCTCTCTGCGCAGCTCATCAGTACCCAG 5340  
Qy |||||  
Db 5281 GCTAAGCGTAGGCTGCGCAGGGGATCTCCCGCTCTCTGCGCAGCTCATCAGTACCCAG 5340  
Qy |||||  
Db 5341 CTGCTGCGCTTCTTGAAGGCAACATGACCTACCCGTCATGACTCCCGGAGCTGAC 5400  
Qy |||||

Db 5341 CTGCTGCGCTTCTCTGAAAGGCAACATGCACTACCCGTCATGACTCCCGGACCGTGCAC 5400  
Qy |||||  
Db 5401 CTCTATCGAGGCGCAACCTCTCTGTGGCGGAGGAGATGGGCGGCAACATACCCCGCTGGAG 5460  
Qy |||||  
Db 5401 CTCTATCGAGGCGCAACCTCTCTGTGGCGGAGGAGATGGGCGGCAACATACCCCGCTGGAG 5460  
Qy |||||  
Db 5461 TCAGAAAATAGAGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAGCGGAGGATGAG 5520  
Qy |||||  
Db 5461 TCAGAAAATAGAGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAGCGGAGGATGAG 5520  
Qy |||||  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTGCGGAGGTCCAGGAATTCCTTCGAGGATG 5580  
Qy |||||  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTGCGGAGGTCCAGGAATTCCTTCGAGGATG 5580  
Qy |||||  
Db 5581 CCCATATGGGACGCGCCCGGATTAACAACCTCCACTGTTTAGAGTCTCTGGAAGGACCCGAC 5640  
Qy |||||  
Db 5581 CCCATATGGGACGCGCCCGGATTAACAACCTCCACTGTTTAGAGTCTCTGGAAGGACCCGAC 5640  
Qy |||||  
Db 5641 TAGCTCCCTCCAGTGGTACAGGGGTGTCATTTGCGGCTGCCAAGGCCCTCCGATACCA 5700  
Qy |||||  
Db 5641 TAGCTCCCTCCAGTGGTACAGGGGTGTCATTTGCGGCTGCCAAGGCCCTCCGATACCA 5700  
Qy |||||  
Db 5701 CCTCCACGAGAGAGAGACGTTGTCCTGTCCAGATCTACGTTGTCCTGCTTGGG 5760  
Qy |||||  
Db 5701 CCTCCACGAGAGAGAGACGTTGTCCTGTCCAGATCTACGTTGTCCTGCTTGGG 5760  
Qy |||||  
Db 5761 GAGCTCGCCCAAAAAGACCTTCGGCAGCTCCGAATCGTCGGCGCTCGACAGCGGACGCA 5820  
Qy |||||  
Db 5761 GAGCTCGCCCAAAAAGACCTTCGGCAGCTCCGAATCGTCGGCGCTCGACAGCGGACGCA 5820  
Qy |||||  
Db 5821 ACGGCTCTCTGACAGGCTTCGAGAGCGGCTCCGAGCGGCGCTCGACAGCGGACGCA 5880  
Qy |||||  
Db 5821 ACGGCTCTCTGACAGGCTTCGAGAGCGGCTCCGAGCGGCGCTCGACAGCGGACGCA 5880  
Qy |||||  
Db 5881 TCCTCCATGCCCCCTTTGAGGGGAGCGCGGGGATCCCGATCTTCAGCAGCGGCTCTGG 5940  
Qy |||||  
Db 5881 TCCTCCATGCCCCCTTTGAGGGGAGCGCGGGGATCCCGATCTTCAGCAGCGGCTCTGG 5940  
Qy |||||  
Db 5941 TCTACCGTAAGCAGGAGGCTAGTGAGGACGTCTCTGTGCTCGATGTCCTACACATGG 6000  
Qy |||||  
Db 5941 TCTACCGTAAGCAGGAGGCTAGTGAGGACGTCTCTGTGCTCGATGTCCTACACATGG 6000  
Qy |||||  
Db 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Qy |||||  
Db 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Qy |||||  
Db 6061 AGCAACTCTTTGCTCCGTCACCACTTGGTCTATGCTACCACTCTCGAGCGCAAGC 6120  
Qy |||||  
Db 6061 AGCAACTCTTTGCTCCGTCACCACTTGGTCTATGCTACCACTCTCGAGCGCAAGC 6120  
Qy |||||  
Db 6121 CTGCGGCAAGAGGTCACCTTTGACAGACTGCGAGTCTCGAGGCTCGGACCACTACCGGAC 6180  
Qy |||||  
Db 6121 CTGCGGCAAGAGGTCACCTTTGACAGACTGCGAGTCTCGAGGCTCGGACCACTACCGGAC 6180  
Qy |||||  
Db 6181 GTGCTCAAGAGATGAAGGCGTAAGCGTCCACAGTTAAGGCTAAACTTCTATCCGTTGAG 6240  
Qy |||||  
Db 6181 GTGCTCAAGAGATGAAGGCGTAAGCGTCCACAGTTAAGGCTAAACTTCTATCCGTTGAG 6240  
Qy |||||  
Db 6241 GAAGCTGTAAAGTGAAGGCGGCGGACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
Qy |||||  
Db 6241 GAAGCTGTAAAGTGAAGGCGGCGGACATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
Qy |||||  
Db 6301 GAGCTCCGGAACCTATTCAGAGGCGGTTAAACCATCCGCTCCGTGTGGAGGAGGACTTG 6360  
Qy |||||  
Db 6301 GAGCTCCGGAACCTATTCAGAGGCGGTTAAACCATCCGCTCCGTGTGGAGGAGGACTTG 6360  
Qy |||||  
Db 6361 CTGGAAGACACTGAGACCAATTCACACCACTCATGCGCAAAATGAGTTTCTGC 6420  
Qy |||||  
Db 6361 CTGGAAGACACTGAGACCAATTCACACCACTCATGCGCAAAATGAGTTTCTGC 6420  
Qy |||||  
Db 6421 GTCCACACAGAGAGGGGCGGAGCGCTTATCGTATTCGATTCAGATTGGGG 6480  
Qy |||||  
Db 6421 GTCCACACAGAGAGGGGCGGAGCGCTTATCGTATTCGATTCAGATTGGGG 6480

QY 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGCCGTG 6540  
 Db 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGCCGTG 6540  
 QY 6541 ATGGGCTCTTCATACGGATTCGATACCTCTCTGACAGCGGCTCGAGTTCCTGTGAAT 6600  
 Db 6541 ATGGGCTCTTCATACGGATTCGATACCTCTCTGACAGCGGCTCGAGTTCCTGTGAAT 6600  
 QY 6601 GCTCGAAGCGAAGAAATGGCCCTTTGCGATATGACACCCCGTGTGTTGACTCA 6660  
 Db 6601 GCTCGAAGCGAAGAAATGGCCCTTTGCGATATGACACCCCGTGTGTTGACTCA 6660  
 QY 6661 ACAGTCACTGAGAAATGACATCGGCTTTGAGAGTCAATCTACCAATGTTGACTGGCC 6720  
 Db 6661 ACAGTCACTGAGAAATGACATCGGCTTTGAGAGTCAATCTACCAATGTTGACTGGCC 6720  
 QY 6721 CCGGAGCCAGACAGCCATAGGCTCGCTACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
 Db 6721 CCGGAGCCAGACAGCCATAGGCTCGCTACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
 QY 6781 ACTAATTTCTAAGGGCAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGACTGAGC 6840  
 Db 6781 ACTAATTTCTAAGGGCAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGACTGAGC 6840  
 QY 6841 ACCAGTGGGTAATACCTCTACATGTTACTTGAAGCGCGCTGCGCGCTGCGAGTGGG 6900  
 Db 6841 ACCAGTGGGTAATACCTCTACATGTTACTTGAAGCGCGCTGCGCGCTGCGAGTGGG 6900  
 QY 6901 AAGCTCAGGACTGACGATGCTGATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
 Db 6901 AAGCTCAGGACTGACGATGCTGATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
 QY 6961 GCGGGACCCACAGAGGAGCGAGCGCTACGCGCTTACGAGGCTATGACTAGATAC 7020  
 Db 6961 GCGGGACCCACAGAGGAGCGAGCGCTACGCGCTTACGAGGCTATGACTAGATAC 7020  
 QY 7021 TCTGCCCCCTCGGGGACCGCCCAACACAGATACGACTTGAAGTGAATCAATCATGC 7080  
 Db 7021 TCTGCCCCCTCGGGGACCGCCCAACACAGATACGACTTGAAGTGAATCAATCATGC 7080  
 QY 7081 TCTTCAATGTGTCAGTGGCGAGTGCATCTGGCAAAAGGCTGACTATCTCACCCGT 7140  
 Db 7081 TCTTCAATGTGTCAGTGGCGAGTGCATCTGGCAAAAGGCTGACTATCTCACCCGT 7140  
 QY 7141 GACCCACACCCCTTTCGCGGCTGCGTGGGAGACAGTACACATCCAGTCAAT 7200  
 Db 7141 GACCCACACCCCTTTCGCGGCTGCGTGGGAGACAGTACACATCCAGTCAAT 7200  
 QY 7201 TCTTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGCAAGGATGATCCTGATG 7260  
 Db 7201 TCTTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGCAAGGATGATCCTGATG 7260  
 QY 7261 ACTCATTTCTTCCATCTCTAGCTCAGGAACAACTTGAAAGGCCCTAGATTGTGAG 7320  
 Db 7261 ACTCATTTCTTCCATCTCTAGCTCAGGAACAACTTGAAAGGCCCTAGATTGTGAG 7320  
 QY 7321 ATCTACGGGCGCTTACTTCAATGAGCACTTACCTTACCTCAGATCATTCACAGCTC 7380  
 Db 7321 ATCTACGGGCGCTTACTTCAATGAGCACTTACCTTACCTCAGATCATTCACAGCTC 7380  
 QY 7381 CATGGCTTAGCGATTTTCACTCATAGTTACTTCTCCAGTGTAGATCAATAGGTTGCT 7440  
 Db 7381 CATGGCTTAGCGATTTTCACTCATAGTTACTTCTCCAGTGTAGATCAATAGGTTGCT 7440  
 QY 7441 TCATGCTTCAGAAACTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGT 7500  
 Db 7441 TCATGCTTCAGAAACTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGT 7500  
 QY 7501 GTCCGCGTACTGCTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
 Db 7501 GTCCGCGTACTGCTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560

QY 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620  
 Db 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620  
 QY 7621 TTATCCAGCTGGTTCGTTGCTGTACAGCGGGGAGACATATATCAAGAGCTGTCTCGT 7680  
 Db 7621 TTATCCAGCTGGTTCGTTGCTGTACAGCGGGGAGACATATATCAAGAGCTGTCTCGT 7680  
 QY 7681 GCCGACCCCGCTGCTTCAATGTGTGCTTACTTCTTCTAGGGGTAGGCACTAT 7740  
 Db 7681 GCCGACCCCGCTGCTTCAATGTGTGCTTACTTCTTCTAGGGGTAGGCACTAT 7740  
 QY 7741 CTACTCCCAACCCGATGAACCGGGAGCTAAACACTCAAGGCAATAGGCACTCTCTTT 7800  
 Db 7741 CTACTCCCAACCCGATGAACCGGGAGCTAAACACTCAAGGCAATAGGCACTCTCTTT 7800  
 QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 QY 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7920  
 Db 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7920  
 QY 7921 TAGCTGTGAAGGTCGTCGAGCGCTTGAATGCTGAGAGAGTGTGATCTCTCTCTCT 7980  
 Db 7921 TAGCTGTGAAGGTCGTCGAGCGCTTGAATGCTGAGAGAGTGTGATCTCTCTCTCT 7980  
 QY 7981 AGATCAAGT 7989  
 Db 7981 AGATCAAGT 7989

RESULT 9  
 US-10-005-469-2  
 ; Sequence 2, Application US/10005469  
 ; Publication No. US20020155133A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANADYS Pharmaceuticals, Inc.  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE  
 ; FILE REFERENCES: 0342/1H395US1  
 ; CURRENT APPLICATION NUMBER: US/10/005,469  
 ; PRIOR FILING DATE: 2002-04-18  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV Replicon RNA from cell line HCVR2  
 US-10-005-469-2

Query Match 99.9%; Score 7982.6; DB 13; Length 7992;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7985; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTTGGGGGAGCACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Db 1 GCCAGCCCCGATTTGGGGGAGCACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 QY 61 TCTTCAGCGAAGAGCTGTAGCCATGCGTGTAGTGTGCTGCGAGCTCCAGGAC 120  
 Db 61 TCTTCAGCGAAGAGCTGTAGCCATGCGTGTAGTGTGCTGCGAGCTCCAGGAC 120  
 QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGAAATGCCAG 180  
 Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGAAATGCCAG 180  
 QY 181 GACGACCGGGTCTTTTCTTGGATCAACCCGCTCAATGCTTGAGATTTGGGCGTCCCCC 240



Db 181 GACGACCGGTCCTTTCTTGGATCAACCGCTCAATGCCCTGAGATTGGGCGTCCCCC 240  
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTGTGTACTGCTCTGATAGG 300  
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTGTGTACTGCTCTGATAGG 300  
QY 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
Db 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
QY 361 CTCAAAGAAAACAAAGGGCGCCCATGATGAAACAAGATGGAATTGCAACAGATGCTTC 420  
Db 361 CTCAAAGAAAACAAAGGGCGCCCATGATGAAACAAGATGGAATTGCAACAGATGCTTC 420  
QY 421 CGGCCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATTCGGCTGCT 480  
Db 421 CGGCCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATTCGGCTGCT 480  
QY 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCCGCTTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCCGCTTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCCGGTCCCTGAATGAACTGCAGGACGAGGCAAGCGGCTATCGTGGCTGGCCA 600  
Db 541 ACCTGTCCGGTCCCTGAATGAACTGCAGGACGAGGCAAGCGGCTATCGTGGCTGGCCA 600  
QY 601 CGACGGGTTCCTTGGCAGCTGTCTCGACGTGTCTACGAGCGGGAAGGACTGGC 660  
Db 601 CGACGGGTTCCTTGGCAGCTGTCTCGACGTGTCTACGAGCGGGAAGGACTGGC 660  
QY 661 TGTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGATCTCACTTCCTCTCCCGGAGA 720  
Db 661 TGTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGATCTCACTTCCTCTCCCGGAGA 720  
QY 721 AAGTATCCATCATGGCTGATGCAATGCGGCGCTGCATAGCTTGATCCGGCTACCTGCC 780  
Db 721 AAGTATCCATCATGGCTGATGCAATGCGGCGCTGCATAGCTTGATCCGGCTACCTGCC 780  
QY 781 CATTCGACCAACAGCAACATCGCATCGACGAGCAGCTACTCGGATGGAAGCGGTC 840  
Db 781 CATTCGACCAACAGCAACATCGCATCGACGAGCAGCTACTCGGATGGAAGCGGTC 840  
QY 841 TTGTGATCAGGATGATCTGACGAAGAGCATCAGGGGCTCGGCCAGCGCACTGTTCG 900  
Db 841 TTGTGATCAGGATGATCTGACGAAGAGCATCAGGGGCTCGGCCAGCGCACTGTTCG 900  
QY 901 CAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTCTGACCCATGGCGATGCT 960  
Db 901 CAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTCTGACCCATGGCGATGCT 960  
QY 961 GCTTGGCGAATCATGCTGGAATAGGCGCTTTCTTGGATTATCGAATGCGATGGCGGC 1020  
Db 961 GCTTGGCGAATCATGCTGGAATAGGCGCTTTCTTGGATTATCGAATGCGATGGCGGC 1020  
QY 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTGTGGCTACCGTGATTCGATTCGATTCG 1080  
Db 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTGTGGCTACCGTGATTCGATTCGATTCG 1080  
QY 1081 TTGGCGGGAATGGGTGACCGCTTCTCGTCTTACGGTATCGCGCTCCCGATTCG 1140  
Db 1081 TTGGCGGGAATGGGTGACCGCTTCTCGTCTTACGGTATCGCGCTCCCGATTCG 1140  
QY 1141 AGCGCATCGCTTCTATCGCTTCTTGAACAGTTCCTTCTGAGTTTAAACAGACCAACG 1200  
Db 1141 AGCGCATCGCTTCTATCGCTTCTTGAACAGTTCCTTCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCCCTTAACTGCTGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCCCTTAACTGCTGGC 1260  
QY 1261 CGAAGCGCTTGAATGAAGCGGCTGTGGTGTATATATGTTATTTCCACCATATTG 1320

Db 1261 CGAAGCGCTTGAATGAAGCGGCTGTGGTGTATATATGTTATTTTCCACCATATTG 1320  
QY 1321 CGGTCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTTCTTGCAGGACATTCCT 1380  
Db 1321 CGGTCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTTCTTGCAGGACATTCCT 1380  
QY 1381 AGGGGTCTTTCCCTCTCGCAAGGAATGCAAGGTCTGTGAAATGTCTGTAAGGAAGCA 1440  
Db 1381 AGGGGTCTTTCCCTCTCGCAAGGAATGCAAGGTCTGTGAAATGTCTGTAAGGAAGCA 1440  
QY 1441 GTTCTCTGGAAGCTTCTTGAAGCAAAACGCTGTAGCGACCCCTTGTGAGGCGAGGG 1500  
Db 1441 GTTCTCTGGAAGCTTCTTGAAGCAAAACGCTGTAGCGACCCCTTGTGAGGCGAGGG 1500  
QY 1501 AACCCCCACCTGCGCACAGGTGCTCTGCGGCCAAAAGCCACGTTATTAAGATACACT 1560  
Db 1501 AACCCCCACCTGCGCACAGGTGCTCTGCGGCCAAAAGCCACGTTATTAAGATACACT 1560  
QY 1561 GCAAGGCGGCACAAACCCAGTGCACGTTGTGAGTTGGAATGAGTCAAA 1620  
Db 1561 GCAAGGCGGCACAAACCCAGTGCACGTTGTGAGTTGGAATGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAGGCGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680  
Db 1621 TGGCTCTCTCAAGCGTATTCAACAGGCGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAA 1740  
Db 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAA 1740  
QY 1741 AACGTCTAGGCGCCCGGAAACACGCGGACGTTGTTTCTTTGAAAACACGATAATACC 1800  
Db 1741 AACGTCTAGGCGCCCGGAAACACGCGGACGTTGTTTCTTTGAAAACACGATAATACC 1800  
QY 1801 ATGGCGCTATTACGCGCTACTCCACAGACGCGGCGCTACTTTGCTGCTGATCATCAT 1860  
Db 1801 ATGGCGCTATTACGCGCTACTCCACAGACGCGGCGCTACTTTGCTGCTGATCATCAT 1860  
QY 1861 AGCCTCAAGCGCGGACAGGAACACAGGTGCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Db 1861 AGCCTCAAGCGCGGACAGGAACACAGGTGCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 1980  
Db 1921 ACACATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 1980  
QY 1981 GGTCTAAAGACCTTGGCGGCCAAAGGGGCCAATCACCCAAATGTACACCAATGTGAC 2040  
Db 1981 GGTCTAAAGACCTTGGCGGCCAAAGGGGCCAATCACCCAAATGTACACCAATGTGAC 2040  
QY 2041 CAGGACCTCTGCTGGCTGGCAAGCGCCCGCGGCGGTTCTTGTGACCATGCACCTGC 2100  
Db 2041 CAGGACCTCTGCTGGCTGGCAAGCGCCCGCGGCGGTTCTTGTGACCATGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGTGTGACGATGCGATGCTCATTCGCGTGGCGCGCG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGTGTGACGATGCGATGCTCATTCGCGTGGCGCGCG 2160  
QY 2161 GCGGACAGCAGGCGGAGCTACTCTCCCGAGCGCTCTCTACTTTGAAGGCTCTTCG 2220  
Db 2161 GCGGACAGCAGGCGGAGCTACTCTCCCGAGCGCTCTCTACTTTGAAGGCTCTTCG 2220  
QY 2221 GCGGTCTCACTGTCTGCGCCCTCGGGGCGACGCTGTGGGCTCTTTCGGGCTGGCGTGC 2280  
Db 2221 GCGGTCTCACTGTCTGCGCCCTCGGGGCGACGCTGTGGGCTCTTTCGGGCTGGCGTGC 2280  
QY 2281 ACCGAGGGGTTCGGAAGCGGTGACTTTGTACCGTCTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTCGGAAGCGGTGACTTTGTACCGTCTGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCACGACAACTCGTCCCTCGGGCGGTACCGCAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCACGACAACTCGTCCCTCGGGCGGTACCGCAGACATTCAGGTG 2400



Qy	2401	GCCCATCTACAGCCGCCCTACTGTTAGCGGCAAGACGACTAAGGTGCGGCTCGTATGCA	2460
Db	2401	GCCCATCTACAGCCGCCCTACTGTTAGCGGCAAGACGACTAAGGTGCGGCTCGTATGCA	2460
Qy	2461	GCCCAAGGGTATAGGTGCTTGTCTCAACCCGTCGCTCGCGCCACCCCTAGTGTTCGGG	2520
Db	2461	GCCCAAGGGTATAGGTGCTTGTCTCAACCCGTCGCTCGCGCCACCCCTAGTGTTCGGG	2520
Qy	2521	GCGTATATGCTTAAGGCACATGTTATCGACCCCTAACATCAGAACCCGGGTAAAGCAATC	2580
Db	2521	GCGTATGCTCTAAGGCACATGTTATCGACCCCTAACATCAGAACCCGGGTAAAGCAATC	2580
Qy	2581	ACCAAGGGTGCCTCCATCAGTACTCCCACTATGGCAAGTTCCTGCGGACGGTGTGTC	2640
Db	2581	ACCAAGGGTGCCTCCATCAGTACTCCCACTATGGCAAGTTCCTGCGGACGGTGTGTC	2640
Qy	2641	TCTGGGGCGCTATGATCATATATGATGAGTGCCACTCAACTGACTCGACCACT	2700
Db	2641	TCTGGGGCGCTATGATCATATATGATGAGTGCCACTCAACTGACTCGACCACT	2700
Qy	2701	ATCCTGGGCATCGGCACAGTCTCGACCAAGCGGAGCGGTCGAGCGCACTCGTGGT	2760
Db	2701	ATCCTGGGCATCGGCACAGTCTCGACCAAGCGGAGCGGTCGAGCGCACTCGTGGT	2760
Qy	2761	CTGCGCACCGCTAGCCCTCCGGGATCGGTGACCCGTGCGCATCCAAACATCGAGGAGGT	2820
Db	2761	CTGCGCACCGCTAGCCCTCCGGGATCGGTGACCCGTGCGCATCCAAACATCGAGGAGGT	2820
Qy	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGGCAATCCCATCGAGACCACT	2880
Db	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGGCAATCCCATCGAGACCACT	2880
Qy	2881	AAGGGGGGAGGCACTCAITTTCTGCGCAATCCAGAAAGAAATGTGATGAGCTCGCGCG	2940
Db	2881	AAGGGGGGAGGCACTCAITTTCTGCGCAATCCAGAAAGAAATGTGATGAGCTCGCGCG	2940
Qy	2941	AAGCTGTCCGGCTCGCATCAATGCTGTAGCATATACCGGGGGCTTGATGATCCGTC	3000
Db	2941	AAGCTGTCCGGCTCGCATCAATGCTGTAGCATATACCGGGGGCTTGATGATCCGTC	3000
Qy	3001	ATACCAACTAGCGGAGACGTCAATGCTGTAGCAAGCGGCTCTAATGACGGGCTTTACC	3060
Db	3001	ATACCAACTAGCGGAGACGTCAATGCTGTAGCAAGCGGCTCTAATGACGGGCTTTACC	3060
Qy	3061	GGCGATTTGCACTCAGTGTATCGACTGCAATACANGTGTCAACCGAGACAGTGCATTCAGC	3120
Db	3061	GGCGATTTGCACTCAGTGTATCGACTGCAATACANGTGTCAACCGAGACAGTGCATTCAGC	3120
Qy	3121	CTGGACCCGACCTTCAACATTTAGACGACGACCGGTGCCAAGACGCGGTGTCAGCTCG	3180
Db	3121	CTGGACCCGACCTTCAACATTTAGACGACGACCGGTGCCAAGACGCGGTGTCAGCTCG	3180
Qy	3181	CAGCGGCGAGGACGACTGTTAGGAGGAGGATGAGGCAATTTACAGTCTTGTGACTCCAGGA	3240
Db	3181	CAGCGGCGAGGACGACTGTTAGGAGGAGGATGAGGCAATTTACAGTCTTGTGACTCCAGGA	3240
Qy	3241	GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTATGACCGGGCTGT	3300
Db	3241	GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTATGACCGGGCTGT	3300
Qy	3301	GCTTGTGATCAGACTCAACCGCCCGGAGACTCAGTTAGTGTTCGGGCTTACCTAACAACA	3360
Db	3301	GCTTGTGATCAGACTCAACCGCCCGGAGACTCAGTTAGTGTTCGGGCTTACCTAACAACA	3360
Qy	3361	CCAGGGTGTCCGCTCTGCCAGGACCACTCTGAGTTCGGGAGAGGCTTTTACAGGCGCTC	3420
Db	3361	CCAGGGTGTCCGCTCTGCCAGGACCACTCTGAGTTCGGGAGAGGCTTTTACAGGCGCTC	3420
Qy	3421	ACCCACATAGACGCCCATTTCTTGTCCGACATTAAGCAGGAGGAGCAACTTCCCGCTAC	3480
Db	3421	ACCCACATAGACGCCCATTTCTTGTCCGACATTAAGCAGGAGGAGCAACTTCCCGCTAC	3480

Db 4561 CAGATCCCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGTATCAAC 4620  
QY 4621 GAGGACTGCTCCACGCGCATGCTCCGCTCGTGGCTAAGAGATGTTTGGATTTGATATGC 4680  
Db 4621 GAGGACTGCTCCACGCGCATGCTCCGCTCGTGGCTAAGAGATGTTTGGATTTGATATGC 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTCGGTCCAGTCCCAAGCTCTCCGCGGATTTGCCGGA 4740  
Db 4681 ACGGTGTTGACTGATTTCAAGACCTCGGTCCAGTCCCAAGCTCTCCGCGGATTTGCCGGA 4740  
QY 4741 GTCCCTTCTCTCATGTCAAGTGGGTCAAGGAGTCTGGCGGGCGACGGCATCATG 4800  
Db 4741 GTCCCTTCTCTCATGTCAAGTGGGTCAAGGAGTCTGGCGGGCGACGGCATCATG 4800  
QY 4801 CAAACACCTGCCCATGTGGAGCACAGATCACCGACATGTGAATAAAGGTTCCATGAGG 4860  
Db 4801 CAAACACCTGCCCATGTGGAGCACAGATCACCGACATGTGAATAAAGGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTCGCCATTAACGCTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTCGCCATTAACGCTAC 4920  
QY 4921 ACCAGGGCCCTGCAAGCCCTCCCGGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG 4980  
Db 4921 ACCAGGGCCCTGCAAGCCCTCCCGGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACATAGTACGGGCGATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACATAGTACGGGCGATG 5040  
QY 5041 ACCACTGACAAAGTAAAGTGCCTGTCAGGTTCGGGCCCGCAATTCCTCACAGAGTG 5100  
Db 5041 ACCACTGACAAAGTAAAGTGCCTGTCAGGTTCGGGCCCGCAATTCCTCACAGAGTG 5100  
QY 5101 GATGGGTGGGTGGTGCACAGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGTC 5160  
Db 5101 GATGGGTGGGTGGTGCACAGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGTC 5160  
QY 5161 ACATTCCTGTCGGGCTCAATCAATCACTGTTGGGTCAAGCTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTCCTGTCGGGCTCAATCAATCACTGTTGGGTCAAGCTCCCATGCGAGCCGAA 5220  
QY 5221 CCGGACGTAGCAGTGTCTCCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGAGC 5280  
Db 5221 CCGGACGTAGCAGTGTCTCCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGAGC 5280  
QY 5281 GCTAAGCTTAGGCTGGCGAGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
Db 5281 GCTAAGCTTAGGCTGGCGAGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
QY 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGAGC 5400  
Db 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGAGC 5400  
QY 5401 CTATCGAGCCCAACCTCTGTCGGCGGAGATGCGGCGGGAACATCAACCCGCTGGAG 5460  
Db 5401 CTATCGAGCCCAACCTCTGTCGGCGGAGATGCGGCGGGAACATCAACCCGCTGGAG 5460  
QY 5461 TCAGAAATTAAGTATGTAATTTTGGACTCTTCGAGCCGCTCCAAAGCGGAGGAGTATG 5520  
Db 5461 TCAGAAATTAAGTATGTAATTTTGGACTCTTCGAGCCGCTCCAAAGCGGAGGAGTATG 5520  
QY 5521 AGGGAAGTATCCGTCGCGGAGATCTCGGAGGTTCAGGAAATTCCTCCAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTCGCGGAGATCTCGGAGGTTCAGGAAATTCCTCCAGCGATG 5580  
QY 5581 CCATATGGGCAAGCCGAGATTAACACCTCCACTGTTAGAGTCTCGGAGGAGCCGAG 5640  
Db 5581 CCATATGGGCAAGCCGAGATTAACACCTCCACTGTTAGAGTCTCGGAGGAGCCGAG 5640  
QY 5641 TAGCTCCCTCCAGTGTACAGGCTGTCCATTTCCGCTGCGCAAGGCCCTCCGATACCA 5700  
Db 5641 TAGCTCCCTCCAGTGTACAGGCTGTCCATTTCCGCTGCGCAAGGCCCTCCGATACCA 5700

Db 5641 TAGCTCCCTCCAGTGTACAGGCTGTCCATTTGCCGCTGCCAAAGGCCCTCCGATACCA 5700  
QY 5701 CTTCCACGAGGAAGAGGACGGTTCCTGTCTCAGAAATCTACCGTGTCTTCTGCTTGGCG 5760  
Db 5701 CTTCCACGAGGAAGAGGACGGTTCCTGTCTCAGAAATCTACCGTGTCTTCTGCTTGGCG 5760  
QY 5761 GAGCTCGGCAAAAGACCTTTGGAGGAGCGGGGATCCCGATCTCAGCGACGGTCTTGG 5820  
Db 5761 GAGCTCGGCAAAAGACCTTTGGAGGAGCGGGGATCCCGATCTCAGCGACGGTCTTGG 5820  
QY 5821 ACGGCTCTCTCCTGACCCAGCCCTCCGACGACGCGGAGTCCGAGCTTGAGTCTGATC 5880  
Db 5821 ACGGCTCTCTCCTGACCCAGCCCTCCGACGACGCGGAGTCCGAGCTTGAGTCTGATC 5880  
QY 5881 TCTTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGCTTGG 5940  
Db 5881 TCTTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGCGACGGGCTTGG 5940  
QY 5941 TCTACCGTAAAGGAGGAGGCTTAGTGAGGACGCTGCTGCTGCTGATCTCTACACATG 6000  
Db 5941 TCTACCGTAAAGGAGGAGGCTTAGTGAGGACGCTGCTGCTGCTGATCTCTACACATG 6000  
QY 6001 ACAGCGCCCTGATCACGCCATCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Db 6001 ACAGCGCCCTGATCACGCCATCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCCGTCACCAAACTTGTCTATGCTACAAATCTCGGAGGCAAGC 6120  
Db 6061 AGCAACTCTTTGCTCCGTCACCAAACTTGTCTATGCTACAAATCTCGGAGGCAAGC 6120  
QY 6121 CTGGGCGAGAAAGGCTACCTTTGACAGACTGCGAGTCTCTGGAGCACCCTACCGGAC 6180  
Db 6121 CTGGGCGAGAAAGGCTACCTTTGACAGACTGCGAGTCTCTGGAGCACCCTACCGGAC 6180  
QY 6181 GTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTAACTTCTATCGTGAG 6240  
Db 6181 GTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTAACTTCTATCGTGAG 6240  
QY 6241 GAAGCTGTAAAGCTGAAGCGCCCAATTCGCGCAGATCTAAATTTGGCTATGGGCGAAAG 6300  
Db 6241 GAAGCTGTAAAGCTGAAGCGCCCAATTCGCGCAGATCTAAATTTGGCTATGGGCGAAAG 6300  
QY 6301 GAGCTCGGAAACCTATCCAGAAAGCGCTTAAACCAATCCGCTCGGTGCGAAGGACTG 6360  
Db 6301 GAGCTCGGAAACCTATCCAGAAAGCGCTTAAACCAATCCGCTCGGTGCGAAGGACTG 6360  
QY 6361 CTGGAAGACTGAGACACCAATTTGACACCAATTTGACCAATTTGACCAATTTGAGTTTCTGC 6420  
Db 6361 CTGGAAGACTGAGACACCAATTTGACCAATTTGACCAATTTGAGTTTCTGC 6420  
QY 6421 GTCCAAACGAGAGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGATTTGGGG 6480  
Db 6421 GTCCAAACGAGAGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGATTTGGGG 6480  
QY 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG 6540  
Db 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG 6540  
QY 6541 ATGGGCTCTTTATACGATTTCCAAATCTCTCCTGGACAGCGGTCCAGTCTCTGTGTAAT 6600  
Db 6541 ATGGGCTCTTTATACGATTTCCAAATCTCTCCTGGACAGCGGTCCAGTCTCTGTGTAAT 6600  
QY 6601 GCTTGAAGGCGAAGAAATGCCCTATGGCTTCGATATGACACCGCTGTTTGTACTCA 6660  
Db 6601 GCTTGAAGGCGAAGAAATGCCCTATGGCTTCGATATGACACCGCTGTTTGTACTCA 6660  
QY 6661 ACGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATTTGAGTTGACTTGGCC 6720  
Db 6661 ACGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATTTGAGTTGACTTGGCC 6720  
QY 6721 CCGGAGCCGAGACAGCCATAGGTCGCTACAGAGCGGCTTTACATCGGGGCGCCCTG 6780  
Db 6721 CCGGAGCCGAGACAGCCATAGGTCGCTACAGAGCGGCTTTACATCGGGGCGCCCTG 6780

6781 ACTAATCTAAAGGACAGATGCGGCTATGCGCGTTCGCGGAGCGGCTGATGACG 6840  
 Db  
 6781 ACTAATCTAAAGGACAGATGCGGCTATGCGCGTTCGCGGAGCGGCTGATGACG 6840  
 Qy  
 6841 ACCAGCTGGGTAAATACCTTACATGTTACTTGAAGGCGGCTGCGGCTGTCGAGCTGG 6900  
 Db  
 6841 ACCAGCTGGGTAAATACCTTACATGTTACTTGAAGGCGGCTGCGGCTGTCGAGCTGG 6900  
 Qy  
 6901 AAGCTCCAGGACTGACAGGATGCTGATGCGGAGACGACTTGTGTTATCTGTGAAAGC 6960  
 Db  
 6901 AAGCTCCAGGACTGACAGGATGCTGATGCGGAGACGACTTGTGTTATCTGTGAAAGC 6960  
 Qy  
 6961 GCGGGGACCCAGAGACGAGGCGGCTTACGCGGCTTACGAGGCTTATGATGATAC 7020  
 Db  
 6961 GCGGGGACCCAGAGACGAGGCGGCTTACGCGGCTTACGAGGCTTATGATGATAC 7020  
 Qy  
 7021 TCTGCCCCCTGGGACCGGCGGCTTACGAGGCTTACGAGGCTTATGATGATAC 7080  
 Db  
 7021 TCTGCCCCCTGGGACCGGCGGCTTACGAGGCTTACGAGGCTTATGATGATAC 7080  
 Qy  
 7081 TCTTCAATGCTGCTGCTGCGGACGATGCTTGGGAAAGGCTTACTTCTCACCGT 7140  
 Db  
 7081 TCTTCAATGCTGCTGCTGCGGACGATGCTTGGGAAAGGCTTACTTCTCACCGT 7140  
 Qy  
 7141 GACCCACACCCCTTGGCGGCTGCTGCGGAGACGAGCTAGACACATCAGTCAAT 7200  
 Db  
 7141 GACCCACACCCCTTGGCGGCTGCTGCGGAGACGAGCTAGACACATCAGTCAAT 7200  
 Qy  
 7201 TCTGCTAGGCAACATCATGATGATGCGGCGGCTTGGGCAAGGATGATCCTGATG 7260  
 Db  
 7201 TCTGCTAGGCAACATCATGATGATGCGGCGGCTTGGGCAAGGATGATCCTGATG 7260  
 Qy  
 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGAAGGCTTGAAGGCTTGAAG 7320  
 Db  
 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGAAGGCTTGAAGGCTTGAAG 7320  
 Qy  
 7321 ATCTAGGCGGCTTACTCCATGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7380  
 Db  
 7321 ATCTAGGCGGCTTACTCCATGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7380  
 Qy  
 7381 CATGGCTTACGCGATTTTCACTCCATGATTTCTCCAGGCTGAGATCAATAGGCTGG 7440  
 Db  
 7381 CATGGCTTACGCGATTTTCACTCCATGATTTCTCCAGGCTGAGATCAATAGGCTGG 7440  
 Qy  
 7441 TCATGCTCAGGAACTTGGGCTACCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 7500  
 Db  
 7441 TCATGCTCAGGAACTTGGGCTACCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 7500  
 Qy  
 7501 GTCGCGCTAGGCTTACTGTCGCGGCGGAGGCTGCGCTTGTGGCAAGTACTCTTC 7560  
 Db  
 7501 GTCGCGCTAGGCTTACTGTCGCGGCGGAGGCTGCGCTTGTGGCAAGTACTCTTC 7560  
 Qy  
 7561 AACTGGGCTAGGACCAAGCTCAAACTCACTCCATCCGCGCTGCGCTGCGCTGCGAT 7620  
 Db  
 7561 AACTGGGCTAGGACCAAGCTCAAACTCACTCCATCCGCGCTGCGCTGCGCTGCGAT 7620  
 Qy  
 7621 TTATCCAGCTGGTTCGTTGCTGTTTACGCGGCGGAGACATATATACAGCTTCTCGT 7680  
 Db  
 7621 TTATCCAGCTGGTTCGTTGCTGTTTACGCGGCGGAGACATATATACAGCTTCTCGT 7680  
 Qy  
 7681 GCCGACCCCGCTGTTTATGTTGCTGCTTACTCTTCTTCTGTTAGGCTAGGATCTAT 7740  
 Db  
 7681 GCCGACCCCGCTGTTTATGTTGCTGCTTACTCTTCTTCTGTTAGGCTAGGATCTAT 7740  
 Qy  
 7741 CTACTCCGACCAAGGAGCTTAAACCTTCCAGGCTTAAAGGCTTAAAGGCTTAAAGG 7800  
 Db  
 7741 CTACTCCGACCAAGGAGCTTAAACCTTCCAGGCTTAAAGGCTTAAAGGCTTAAAGG 7800  
 Qy  
 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7860  
 Db  
 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7860

7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Db  
 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Qy  
 7921 TAGCTGTGAAAGGCTGCTGAGCGGCTTACCTGACAGAGTGTGATGATGCTGCTG 7980  
 Db  
 7921 TAGCTGTGAAAGGCTGCTGAGCGGCTTACCTGACAGAGTGTGATGATGCTGCTG 7980  
 Qy  
 7981 AGATCAAGT 7989  
 Db  
 7981 AGATCAAGT 7989

RESULT 10  
 US-10-005-469-5  
 ; Sequence 5, Application US/10005469  
 ; Publication No. US20020155133A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANADYS Pharmaceuticals, Inc.  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REP  
 ; FILE REFERENCE: 0342/1H395US1  
 ; CURRENT APPLICATION NUMBER: US/10/005,469  
 ; CURRENT FILING DATE: 2002-04-18  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV Replicon from cell line HCV22  
 ; US-10-005-469-5

Query Match 99.9%; Score 7982.6; DB 13; Length 7992;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7985; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCCAGCCCCGATTTGGGGGCGACCTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Db 1 GCCAGCCCCGATTTGGGGGCGACCTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Qy 61 TCTTCAGGCAAGGCTCTAGCCATGGCTTAGTATGAGTGTGTCGACCTCCAGGAC 120  
 Db 61 TCTTCAGGCAAGGCTCTAGCCATGGCTTAGTATGAGTGTGTCGACCTCCAGGAC 120  
 Qy 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
 Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
 Qy 181 GAGCAGCGGTCTTTCTTGATCAACCCCTCAATGCTGAGATTTGGCGTGGCCCC 240  
 Db 181 GAGCAGCGGTCTTTCTTGATCAACCCCTCAATGCTGAGATTTGGCGTGGCCCC 240  
 Qy 241 GCGGAGCTGTAGCCGAGTGTGTTGGTTCGAAAGGCTTGTGTTACTGCTGATAGG 300  
 Db 241 GCGGAGCTGTAGCCGAGTGTGTTGGTTCGAAAGGCTTGTGTTACTGCTGATAGG 300  
 Qy 301 GTGCTTCGAGTGTCCCGGAGGCTCTGTAGACCGGTGACCATGAGCACGAATCCTAAAC 360  
 Db 301 GTGCTTCGAGTGTCCCGGAGGCTCTGTAGACCGGTGACCATGAGCACGAATCCTAAAC 360  
 Qy 361 CTCAGAGAAAACCAAGGCGGCCCATGATTTGAACAGATGATTTGACGAGGTTCTC 420  
 Db 361 CTCAGAGAAAACCAAGGCGGCCCATGATTTGAACAGATGATTTGACGAGGTTCTC 420  
 Qy 421 CGGCGCTTGGTGGAGGCTATTTCGCTATGATGCTGGGCAACAGCAATTCGCTGCT 480  
 Db 421 CGGCGCTTGGTGGAGGCTATTTCGCTATGATGCTGGGCAACAGCAATTCGCTGCT 480  
 Qy 481 CTGATGCGCGCTGTTCGGGCTGTACGCGAGGGGCGGCCGGTCTTTTGTCAAGCG 540

Db 481 CTGATCGCGCTGTCGGCTGTACGAGGGCGCCCGTCTTTTGTCAAGACG 540  
QY 541 ACCTGTCCGGTCCCTGAAATGAATCGAGAGAGGAGCGCGCTATCGTGGTGGCCA 600  
Db 541 ACCTGTCCGGTCCCTGAAATGAATCGAGAGAGGAGCGCGCTATCGTGGTGGCCA 600  
QY 601 CGACGGCGTTCCTTGGCGAGCTGTCTCGACGTTGTCAATGAAGCGGGAAGGACTGGC 660  
Db 601 CGACGGCGTTCCTTGGCGAGCTGTCTCGACGTTGTCAATGAAGCGGGAAGGACTGGC 660  
QY 661 TGTATTGGGGAAGTCCGCGGAGGATCTCTGTCTCATCTTGTCTCTCCCGAGA 720  
Db 661 TGTATTGGGGAAGTCCGCGGAGGATCTCTGTCTCATCTCACTTGTCTCTCCCGAGA 720  
QY 721 AAGTATCCATCATTGGCTGATGCAATGCGGCGCTGCAATAGCTTTGATCCCGTACCTGCC 780  
Db 721 AAGTATCCATCATTGGCTGATGCAATGCGGCGCTGCAATAGCTTTGATCCCGTACCTGCC 780  
QY 781 CATTTCGACCAACGAGCAATCGCATCGAGCGAGCACGCTACTCGATGGAGCGCGTC 840  
Db 781 CATTTCGACCAACGAGCAATCGCATCGAGCGAGCACGCTACTCGATGGAGCGCGTC 840  
QY 841 TTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTGC 900  
Db 841 TTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTGC 900  
QY 901 CCAGGCTCAAGGCGCGATCCCGAGCGAGGATCTCGTGGTGAAGGATGATGTCG 960  
Db 901 CCAGGCTCAAGGCGCGATCCCGAGCGAGGATCTCGTGGTGAAGGATGATGTCG 960  
QY 961 GCTTGGCGAATATCATGGTGAATGGCGCTTTCTTGATTCATCGACTGTGGCGGC 1020  
Db 961 GCTTGGCGAATATCATGGTGAATGGCGCTTTCTTGATTCATCGACTGTGGCGGC 1020  
QY 1021 TGGGTGTGGCGACCGCTATCAGACATAGCTGTGGTACCGGTGATTTGCTGAAGAGC 1080  
Db 1021 TGGGTGTGGCGACCGCTATCAGACATAGCTGTGGTACCGGTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGGCAATGGGTGACCGTCTCTGTGCTTTACGTATGCGCGTCCCGATTCG 1140  
Db 1081 TTGGCGGCAATGGGTGACCGTCTCTGTGCTTTACGTATGCGCGTCCCGATTCG 1140  
QY 1141 AGCGATCGCTTCTATCGCTTCTTGACAGTCTCTTGTGAGTTTAAACAGACCAACG 1200  
Db 1141 AGCGATCGCTTCTATCGCTTCTTGACAGTCTCTTGTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCTCAACGTTACTGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCTCAACGTTACTGGC 1260  
QY 1261 CGAAGCGCTTGAATAGAGCGCGTGTGTTGTCTATGTTATTTTCCACCATTTG 1320  
Db 1261 CGAAGCGCTTGAATAGAGCGCGTGTGTTGTCTATGTTATTTTCCACCATTTG 1320  
QY 1321 CGGTCTTTTGGCAATGTAGGGCCCGGAAACCTTGTGATTCGTTGAAGGAGCA 1440  
Db 1321 CGGTCTTTTGGCAATGTAGGGCCCGGAAACCTTGTGATTCGTTGAAGGAGCA 1440  
QY 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATTCGTTGAAGGAGCA 1500  
Db 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATTCGTTGAAGGAGCA 1500  
QY 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGAGGACAGCG 1560  
Db 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTGAGGACAGCG 1560  
QY 1501 AACCCCGACCTTGGCGACAGGTGCTCTGGCGCAAGGCGCGTGTATGAATACACT 1560  
Db 1501 AACCCCGACCTTGGCGACAGGTGCTCTGGCGCAAGGCGCGTGTATGAATACACT 1560  
QY 1561 GCAAGCGGCGCAACCCCGAGTGCACGTTGTGATGATGTTGATGTTGGAAGAGTCAAA 1620  
Db 1561 GCAAGCGGCGCAACCCCGAGTGCACGTTGTGATGATGTTGATGTTGGAAGAGTCAAA 1620

Db 1561 GCAAGCGGCGCAACCCCGAGTGCACGTTGTGATGATGTTGATGTTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAGGTATCCCATTTGT 1680  
Db 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAGGTATCCCATTTGT 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCGGTGACATGCTTTTACATGTTTGTAGTCGAGGTTAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCGGTGACATGCTTTTACATGTTTGTAGTCGAGGTTAAA 1740  
QY 1741 AACGTCTAGGCGCCCGGAAACACAGGGGACGTTTCTTCCCTTGAAGAACAGATAATACC 1800  
Db 1741 AACGTCTAGGCGCCCGGAAACACAGGGGACGTTTCTTCCCTTGAAGAACAGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCAATCACT 1860  
Db 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCAATCACT 1860  
QY 1861 AGCTCACAGCGCGGACAGAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTCACAGCGCGGACAGAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTCTTGGCGACTCGTCAATGGCGTGTGGTGTGAGTGTCTATCATGTTGCC 1980  
Db 1921 ACACATCTTCTTGGCGACTCGTCAATGGCGTGTGGTGTGAGTGTCTATCATGTTGCC 1980  
QY 1981 GGCTCAAGACCTTGTGCGGCCCAAGGGCCCAATCAACCAATGTACCAATGTGAGC 2040  
Db 1981 GGCTCAAGACCTTGTGCGGCCCAAGGGCCCAATCAACCAATGTACCAATGTGAGC 2040  
QY 2041 CAGGACTCGTGGCTGGGAGCGCCCGCGGGCGGCTTCTTGTGACCAATGCACTGC 2100  
Db 2041 CAGGACTCGTGGCTGGGAGCGCCCGCGGGCGGCTTCTTGTGACCAATGCACTGC 2100  
QY 2101 GGAGCTCGGACTTGTGAGTCAAGAGCATGCGATGTCTTCTTCTTCTTCTTCTTCTTCT 2160  
Db 2101 GGAGCTCGGACTTGTGAGTCAAGAGCATGCGATGTCTTCTTCTTCTTCTTCTTCTTCT 2160  
QY 2161 GGCGACAGAGGGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
Db 2161 GGCGACAGAGGGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2220  
QY 2221 GGCGGTCATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2280  
Db 2221 GGCGGTCATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2280  
QY 2281 ACCGAGGGGTTGCGAAGCGGTGAGTCTTGTACCGTGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTGCGAAGCGGTGAGTCTTGTACCGTGTGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTCAACGAGCAAACTGCTCCCTTCCCGCGGTACCGCAGACATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTCAACGAGCAAACTGCTCCCTTCCCGCGGTACCGCAGACATTTCCAGGTG 2400  
QY 2401 GCGCATCTACAGCCCTTCTGCTAGCGGCAAGAGCATTAAGTGTGCGGTGCGTATGCA 2460  
Db 2401 GCGCATCTACAGCCCTTCTGCTAGCGGCAAGAGCATTAAGTGTGCGGTGCGTATGCA 2460  
QY 2461 GCGCAAGGGTATAAGGTGCTTGTGACCGGTGCTGCGCGCGCACCTAGTGTTCGGG 2520  
Db 2461 GCGCAAGGGTATAAGGTGCTTGTGACCGGTGCTGCGCGCGCACCTAGTGTTCGGG 2520  
QY 2521 GCGTATATGTTTAAAGGACATGTTATCGACCCCTTAAACATCAGAACCGGGTAAAGCACTC 2580  
Db 2521 GCGTATATGTTTAAAGGACATGTTATCGACCCCTTAAACATCAGAACCGGGTAAAGCACTC 2580  
QY 2581 ACCAGGGTGGCCCATCAGTCTACCTTATGGAAGTGTTCGCGACCGTGTGTC 2640  
Db 2581 ACCAGGGTGGCCCATCAGTCTACCTTATGGAAGTGTTCGCGACCGTGTGTC 2640  
QY 2641 TCTGGGGGCGCTATGACATCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
Db 2641 TCTGGGGGCGCTATGACATCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 2700



Db	4861	ATCTGTGGGCTTAGGACTGTAGTAACACAGTGGGATGAACAATTTCCCAATTAAACGCGTAC	4920
QY	4921	ACCAAGGGCCCTGCACGCCCTCCCGCGCCAAATTAATCTAGGCGCTGTGTGCGGGTG	4980
Db	4921	ACCACGGGCCCTGCACGCCCTCCCGCGCCAAATTAATCTAGGCGCTGTGTGCGGGTG	4980
QY	4981	GCTGCTAGGAGTAGTGTGGAGTTACGGGGTGGGGATTTCCACTACGTGACGGGATG	5040
Db	4981	GCTGCTAGGAGTAGTGTGGAGTTACGGGGTGGGGATTTCCACTACGTGACGGGATG	5040
QY	5041	ACCACTGACAACTGAAAGTGCCCGTGTACAGTTCCGGCCCCCGGAATTTCTCAAGAAAGT	5100
Db	5041	ACCACTGACAACTGAAAGTGCCCGTGTACAGTTCCGGCCCCCGGAATTTCTCAAGAAAGT	5100
QY	5101	GATGGGTTGCGGTTGCACAGTAGTCCTCAGCGTGCAAACCCCTTCCTACGGGAGGAGTTC	5160
Db	5101	GATGGGTTGCGGTTGCACAGTAGTCCTCAGCGTGCAAACCCCTTCCTACGGGAGGAGTTC	5160
QY	5161	ACATTCCTGTGCGGCTCAATCAATACTGTTGGGTACAGTTCCTCAGCGAGGAGTTC	5220
Db	5161	ACATTCCTGTGCGGCTCAATCAATACTGTTGGGTACAGTTCCTCAGCGAGGAGTTC	5220
QY	5221	CCGACAGTGTAGCAGTGTCACTTCATGCTCACGACCCCTCCACATTCACGGCGGAGACG	5280
Db	5221	CCGACAGTGTAGCAGTGTCACTTCATGCTCACGACCCCTCCACATTCACGGCGGAGACG	5280
QY	5281	GCTAAGCGTAGGCTGGCGACGGGATCTCCGCCCTCTTGGCCAGTTCATCAGCTAGCCAG	5340
Db	5281	GCTAAGCGTAGGCTGGCGACGGGATCTCCGCCCTCTTGGCCAGTTCATCAGCTAGCCAG	5340
QY	5341	CTGTCTGGGCTTCTCTTGAAGGCAACATGCACTACCGGTATGACTCCCCGGACGCTGAC	5400
Db	5341	CTGTCTGGGCTTCTCTTGAAGGCAACATGCACTACCGGTATGACTCCCCGGACGCTGAC	5400
QY	5401	CTCATCGAGGCCAACCTCTGTGGCGGACGAGATGGGCGGGAACATCATCCCGCTGGAG	5460
Db	5401	CTCATCGAGGCCAACCTCTGTGGCGGACGAGATGGGCGGGAACATCATCCCGCTGGAG	5460
QY	5461	TCGAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGTATGAG	5520
Db	5461	TCGAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGTATGAG	5520
QY	5521	AGGGAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCCAGGAAATTCCTCGAGCGATG	5580
Db	5521	AGGGAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCCAGGAAATTCCTCGAGCGATG	5580
QY	5581	CCCATATGGCAGCCCGGATTACACCTTCCACTGTAGAGTCTTGAAGGACCCGAC	5640
Db	5581	CCCATATGGCAGCCCGGATTACACCTTCCACTGTAGAGTCTTGAAGGACCCGAC	5640
QY	5641	TACGTCCTCCAGTGTGTACACGGGTGCCATTCGCCGCTGCCAAGGCCCTCCGATACCA	5700
Db	5641	TACGTCCTCCAGTGTGTACACGGGTGCCATTCGCCGCTGCCAAGGCCCTCCGATACCA	5700
QY	5701	CCTCCAGGAGGAGGAGGTTGTCCTGTGAGAAATCTACCGTGTCTTCTGCCTTGGCG	5760
Db	5701	CCTCCAGGAGGAGGAGGAGTGTGTCCTGTGAGAAATCTACCGTGTCTTCTGCCTTGGCG	5760
QY	5761	GAGCTGCCACAAGACCTTCGACGCTCCGAATCGTCGGCGTGTGACAGCGGACCGGCA	5820
Db	5761	GAGCTGCCACAAGACCTTCGACGCTCCGAATCGTCGGCGTGTGACAGCGGACCGGCA	5820
QY	5821	ACGGCCCTCTCCTGACACGCGCTCCGACGAGCGGAGTCCGAGTTCGAGTTCGCTAC	5880
Db	5821	ACGGCCCTCTCCTGACACGCGCTCCGACGAGCGGAGTCCGAGTTCGAGTTCGCTAC	5880
QY	5881	TCCTCATGCCCCCTTGTAGGGGAGCCGGGGATTCGAGTCTCAGCGACGGGCTTTGG	5940
Db	5881	TCCTCATGCCCCCTTGTAGGGGAGCCGGGGATTCGAGTCTCAGCGACGGGCTTTGG	5940
QY	5941	TCCTACCGTAGCGAGGCTGTAGGAGCTGTCTGTCTCGATGTCCTACATGTCCTACATGG	6000

5941	TCTACGTAAGCGAGGAGGCTAGTGAGGACGTCGCTGCTGCTCGATGCTCCTACACATGG	6000
6001	ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCGCCATCAATGCACTGG	6060
6001		
6001	ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTGG	6060
6061	AGCAACTCTTTGCTCCGTCACCAACAACCTTGGTCTATGCTCAACAATCTCGCAGCGCAAGC	6120
6061	AGCAACTCTTTGCTCCGTCACCAACAATTGGTCTATGCTACACATCTCGCAGCGCAAGC	6120
6121	CTGCGGCGAAGAAGGTCACCTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
6121	CTGCGGCGAAGAAGGTCACCTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
6181	GTGCTCAAGGAGATGAAGCGCAAGCGCTCCACAGTTAAGGCTAAACTTCTATCGTGGAG	6240
6181	GTGCTCAAGGAGATGAAGCGCAAGCGCTCCACAGTTAAGGCTAAACTTCTATCGTGGAG	6240
6241	GAAGCCTCTAAGCTGACGCCGCCACATTTGGCGCAGATCTAAATTTGGCTATGGGCAAG	6300
6241	GAAGCCTCTAAGCTGACGCCGCCACATTTGGCGCAGATCTAAATTTGGCTATGGGCAAG	6300
6301	GACGTCGGAAACCTTATCCAGCAAGCGCGTTAAACCATCCGCTCCGTTGCGAAGGACTTG	6360
6301	GACGTCGGAAACCTTATCCAGCAAGCGCGTTAAACCATCCGCTCCGTTGCGAAGGACTTG	6360
6361	CTGGAAGACCTGAGACACCAATTTGACACCAACCATCATGCGCAAAAATGAGTTTCTGC	6420
6361	CTGGAAGACCTGAGACACCAATTTGACACCAACCATCATGCGCAAAAATGAGTTTCTGC	6420
6421	GTCCAAACAGAGAAGGGGGCGCAAGCCAGCTCGCCTTATCTGTTATTTCCACAGATTTGGGG	6480
6421	GTCCAAACAGAGAAGGGGGCGCAAGCCAGCTCGCCTTATCTGTTATTTCCACAGATTTGGGG	6480
6481	GTTCTGTGTGCGAGAAATATGGCCCTTTACGATGTGTTCTCCACCCTCCCTCAGGCCGTG	6540
6481	GTTCTGTGTGCGAGAAATATGGCCCTTTACGATGTGTTCTCCACCCTCCCTCAGGCCGTG	6540
6541	ATGGGCTCTTCATACGATTCCCAATCTCTCTGACGCGGTGCAAGTCTCTGTGTGAAT	6600
6541	ATGGGCTCTTCATACGATTCCCAATCTCTCTGACGCGGTGCAAGTCTCTGTGTGAAT	6600
6601	GCCTGGAAGCGAGAAATCCCTATGGGCTTCGCGATATGACACCCGCTGTTTGTACTCA	6660
6601	GCCTGGAAGCGAGAAATCCCTATGGGCTTCGCGATATGACACCCGCTGTTTGTACTCA	6660
6661	ACGCTCACTAGAAATGACATCCGTTGTGAGGATCAATCTACCAATGTTGTGACTTGGCC	6720
6661	ACGCTCACTAGAAATGACATCCGTTGTGAGGATCAATCTACCAATGTTGTGACTTGGCC	6720
6721	CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGGCGCTTTACATCGGGGCCCCCTG	6780
6721	CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGGCGCTTTACATCGGGGCCCCCTG	6780
6781	ACTAATTTCTAAAGGCGAGAACTCGCGCTATCGCGGTCGCGCGAGCGGTCTACTGACG	6840
6781	ACTAATTTCTAAAGGCGAGAACTCGCGCTATCGCGGTCGCGCGAGCGGTCTACTGACG	6840
6841	ACGAGTCGCGTAAATACCTTCATGTTATTGAAGCGCGCTGCGGCTGTGCAAGTCGG	6900
6841	ACGAGTCGCGTAAATACCTTCATGTTATTGAAGCGCGCTGCGGCTGTGCAAGTCGG	6900
6901	AAGCTCCAGGACTGACGATGCTGTTATCGGAGACGACTTGTGTTATCTGTCAAAAGC	6960
6901	AAGCTCCAGGACTGACGATGCTGTTATCGGAGACGACTTGTGTTATCTGTCAAAAGC	6960
6961	AGCTCCAGGACTGACGATGCTGTTATCGGAGACGACTTGTGTTATCTGTCAAAAGC	6960
6961	AGCTCCAGGACTGACGATGCTGTTATCGGAGACGACTTGTGTTATCTGTCAAAAGC	6960
7021	TCGTCGCCCTCTGGGACCGGCCCAAAACAGAAATACGACTTGGAGTTGATTAACATCATGC	7080
7021	TCGTCGCCCTCTGGGACCGGCCCAAAACAGAAATACGACTTGGAGTTGATTAACATCATGC	7080



```

7081 TCCTCCATGTCAGTCGGCAGCATGATCTGGCAAAAGGCTACTATCTCACCGT 7140
7081 TCCTCCATGTCAGTCGGCAGCATGATCTGGCAAAAGGCTACTATCTCACCGT 7140
7141 GACCCACACACCCCTTCGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
7141 GACCCACACACCCCTTCGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
7201 TCCTGGCTAGGCAACATCATGATGCGCCCACTTGTGGGCAAGGATGCTGTGATG 7260
7201 TCCTGGCTAGGCAACATCATGATGCGCCCACTTGTGGGCAAGGATGCTGTGATG 7260
7261 ACTGATTTCTTCATCTCTTACTAGCTCAGGACAACTTGAAGAGCCCTAGATTGTCAG 7320
7261 ACTGATTTCTTCATCTCTTACTAGCTCAGGACAACTTGAAGAGCCCTAGATTGTCAG 7320
7321 ATCTAGGGGCTGTCTTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATCAAGACTC 7380
7321 ATCTAGGGGCTGTCTTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATCAAGACTC 7380
7381 CATGGCTTAGGCAATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGGTGGCT 7440
7381 CATGGCTTAGGCAATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGGTGGCT 7440
7441 TCATGCTCAGGAACTTGGGTACCGCCCTTGCAGTCTGAGACATCGGGCCAGAGT 7500
7441 TCATGCTCAGGAACTTGGGTACCGCCCTTGCAGTCTGAGACATCGGGCCAGAGT 7500
7501 GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
7501 GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
7561 AACTGGGCACTAGGACCAAGCTCAAACTCACTCCAATCCCGCTGCTGCCAGTTGGAT 7620
7561 AACTGGGCACTAGGACCAAGCTCAAACTCACTCCAATCCCGCTGCTGCCAGTTGGAT 7620
7621 TTATCCAGCTGGTTCGTGTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680
7621 TTATCCAGCTGGTTCGTGTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680
7681 GCGCGACCCGCTGTTTCATGTGCTACTCTACTTCTCTAGGGGTAGGCATCTAT 7740
7681 GCGCGACCCGCTGTTTCATGTGCTACTCTACTTCTCTAGGGGTAGGCATCTAT 7740
7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
7921 TAGCTGTGAAGGTTCGTGAGCCGCTTGAATGAGAGAGTGTGATGACTGGCTCTCTGC 7980
7921 TAGCTGTGAAGGTTCGTGAGCCGCTTGAATGAGAGAGTGTGATGACTGGCTCTCTGC 7980
7981 AGATCAAGT 7989
7981 AGATCAAGT 7989

```

RESULT 11  
 US-10-434-842-2  
 ; Sequence 2, Application US/10434842  
 ; Publication No. US2004005549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REH

```

; FILE REFERENCE: 0342/1H395US3
; CURRENT APPLICATION NUMBER: US/10/434,842
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/233,307
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 10/005,469
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV2 subgenomic HCV replicon
; US-10-434-842-2

Query Match 99.9%; Score 7982.6; DB 16; Length 7992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 7985; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAAGTCTG 60
DB 1 GCCAGCCCCGATTTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAAGTCTG 60
QY 61 TCTTACGCGAGAAAGGCTCTAGCCATGGCTTATGATGATGATGATGATGATGATGATG 120
DB 61 TCTTACGCGAGAAAGGCTCTAGCCATGGCTTATGATGATGATGATGATGATGATGATG 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCG 180
QY 181 GACGACCGGGTCTTTCTTTGGATCAACCGCTCAATGCTTGGAGATTTGGGCTGCCCC 240
DB 181 GACGACCGGGTCTTTCTTTGGATCAACCGCTCAATGCTTGGAGATTTGGGCTGCCCC 240
QY 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
DB 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
QY 301 GTGCTTTCGAGTGCCTCCGGAGGTCTCTAGACCGGTGACCATGAGCAAGATTCCTAAC 360
DB 301 GTGCTTTCGAGTGCCTCCGGAGGTCTCTCTAGACCGGTGACCATGAGCAAGATTCCTAAC 360
QY 361 CTCAAGAAAAACCAAGGGCGGCGCATGATTGAAACAGATGATGATGATGATGATGATGATG 420
DB 361 CTCAAGAAAAACCAAGGGCGGCGCATGATTGAAACAGATGATGATGATGATGATGATG 420
QY 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTGGGCAACAGACAAATCGGCTGCT 480
DB 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTATGCTGGGCAACAGACAAATCGGCTGCT 480
QY 481 CTGATGCGCGCTGTTTCCGGCTCTCAGCGCAGGGGCGCGGTTCTTTTGTCAAGACCG 540
DB 481 CTGATGCGCGCTGTTTCCGGCTCTCAGCGCAGGGGCGCGGTTCTTTTGTCAAGACCG 540
QY 541 ACTGTTCGGGTGCGCTTGAATGAATGAACTGAGGACAGGAGCGCGGCTATCGTGGCTG 600
DB 541 ACTGTTCGGGTGCGCTTGAATGAATGAACTGAGGACAGGAGCGCGGCTATCGTGGCTG 600
QY 601 CGACGGGCTTCTTTCGCGAGCTGTGCTCGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGACGGGCTTCTTTCGCGAGCTGTGCTCGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 TGCTATTGGCGGAAGTGCAGGATCTCTGCTCATCTCACTTGTCTCTGCTCTGCTCTGCTG 720
DB 661 TGCTATTGGCGGAAGTGCAGGATCTCTGCTCATCTCACTTGTCTCTGCTCTGCTCTGCTG 720
QY 721 AAGTATCCATGCTGCTGATGCAATGCGGGCTGATGATGATGATGATGATGATGATGATGATG 780
DB 721 AAGTATCCATGCTGCTGATGCAATGCGGGCTGATGATGATGATGATGATGATGATGATGATG 780

```



Db	721	AAGTATCCATCATGGCTGATGCAATCGCGCGCTGCATACGCTTGATCGGCTACTCGC	780
Qy	781	CATTGACACCAAGCAACATCCATCGACGACGACGACTCGATGGAAGCGGTC	840
Db	781	CATTGACACCAAGCAACATCCATCGACGACGACGACTCGATGGAAGCGGTC	840
Qy	841	TTGTGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTGC	900
Db	841	TTGTGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTGC	900
Qy	901	CCAGCTCAAGCGCGCATGCGCGACGCGAGGATCTCGTCTGTGACCCATGCGCATGCT	960
Db	901	CCAGCTCAAGCGCGCATGCGCGACGCGAGGATCTCGTCTGTGACCCATGCGCATGCT	960
Qy	961	GCTTCCCGAATATCATGTGGAAATGCGCGCTTTCTGGATTGATGACGCTGCGCGCG	1020
Db	961	GCTTCCCGAATATCATGTGGAAATGCGCGCTTTCTGGATTGATGACGCTGCGCGCG	1020
Qy	1021	TGGGTGTCGCGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTGCTGAAGAGC	1080
Db	1021	TGGGTGTCGCGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTGCTGAAGAGC	1080
Qy	1081	TTGGCGGGAATGGCTGACCGCTTCTCGTCTTTACGCTATCGCGCTCCCGATTGCG	1140
Db	1081	TTGGCGGGAATGGCTGACCGCTTCTCGTCTTTACGCTATCGCGCTCCCGATTGCG	1140
Qy	1141	AGCGATCGCTTCTATCGCTTCTGACGATGCTTCTGAGTTCTGAGTTTAAACAGACCAACG	1200
Db	1141	AGCGATCGCTTCTATCGCTTCTGACGATGCTTCTGAGTTCTGAGTTTAAACAGACCAACG	1200
Qy	1201	GTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTCAAGTTACTGCG	1260
Db	1201	GTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTCAAGTTACTGCG	1260
Qy	1261	CGAAGCGCTTGGAAATAGCGCGGTGCTGTTGCTATATGTTTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGCTTGGAAATAGCGCGGTGCTGTTGCTATATGTTTATTTTCCACCATATTG	1320
Qy	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAACTGGCCCTGCTTCTTGAACGAGATTCCT	1380
Db	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAACTGGCCCTGCTTCTTGAACGAGATTCCT	1380
Qy	1381	AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGCTGAAGGAAGCA	1440
Db	1381	AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGCTGAAGGAAGCA	1440
Qy	1441	GTTCTCTGGAAGCTTCTGGAAGCAAAACAGTCTGTAGCGACCTTTGACGAGCAGCG	1500
Db	1441	GTTCTCTGGAAGCTTCTGGAAGCAAAACAGTCTGTAGCGACCTTTGACGAGCAGCG	1500
Qy	1501	AACCCCGCCTGCGGACAGTGCCTCTGCGGCCAAAGCCACGCTGATAGATACACCT	1560
Db	1501	AACCCCGCCTGCGGACAGTGCCTCTGCGGCCAAAGCCACGCTGATAGATACACCT	1560
Qy	1561	GCAAGCGGCACAAACCCAGTCCAGTCTGTGAGTTGGATGTTGGAAGAGTCAAA	1620
Db	1561	GCAAGCGGCACAAACCCAGTCCAGTCTGTGAGTTGGATGTTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCCTCAAGCTATTCAACAGGGGCTGAAGGATGCCAGAAAGTACCCCATTTG	1680
Db	1621	TGGCTCTCCTCAAGCTATTCAACAGGGGCTGAAGGATGCCAGAAAGTACCCCATTTG	1680
Qy	1681	ATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTAGTGAAGTTAAA	1740
Db	1681	ATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTAGTGAAGTTAAA	1740
Qy	1741	AACGTCTAGGCCCGCCGAAACCAAGGACGCTGTTTCTTTGAAACACGATATACC	1800
Db	1741	AACGTCTAGGCCCGCCGAAACCAAGGACGCTGTTTCTTTGAAACACGATATACC	1800
Qy	1801	ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGCGCTACTTGGCTGCATCATCT	1860
Db	1801	ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGCGCTACTTGGCTGCATCATCT	1860
Qy	1861	AGCCTCAAGCGCGGACAGGAACAGGTGCGAGGGAGGTCAAAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCAAGCGCGGACAGGAACAGGTGCGAGGGAGGTCAAAGTGTCTCCACCGCA	1920
Qy	1921	ACACAATCTTCTTGGCGACCTGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGCG	1980
Db	1921	ACACAATCTTCTTGGCGACCTGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGCG	1980
Qy	1981	GGCTCAAAAGACCTTGGCGCGCCCAAGGGCCCAATACCCCAAAATGACCAAAATGAGAC	2040
Db	1981	GGCTCAAAAGACCTTGGCGCGCCCAAGGGCCCAATACCCCAAAATGACCAAAATGAGAC	2040
Qy	2041	CAGACCTCTGTCGGCTGCAAGCGCCCGCGGGCGCTTCTTGAACCAATGACCTGCG	2100
Db	2041	CAGACCTCTGTCGGCTGCAAGCGCCCGCGGGCGCTTCTTGAACCAATGACCTGCG	2100
Qy	2101	GCGAGCTCGGACCTTACTTGGTCAAGGATGCGGATGTCATTCGGTGGCGCGGG	2160
Db	2101	GCGAGCTCGGACCTTACTTGGTCAAGGATGCGGATGTCATTCGGTGGCGCGGG	2160
Qy	2161	GGGACAGCAGCGGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220
Db	2161	GGGACAGCAGCGGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220
Qy	2221	GGCGTCCACTGCTCTGCGGCGCAGCTGCGGCGCTCTTTCGGGCTCTTTCGGGCTGCG	2280
Db	2221	GGCGTCCACTGCTCTGCGGCGCAGCTGCGGCGCTCTTTCGGGCTCTTTCGGGCTGCG	2280
Qy	2281	ACCGAGGGGTTGCGAGGGGCTGACTTGTACCCGTCGAGTCTATGGAACCACTATG	2340
Db	2281	ACCGAGGGGTTGCGAGGGGCTGACTTGTACCCGTCGAGTCTATGGAACCACTATG	2340
Qy	2341	CGGTCCCGGCTTCTCAAGCAAACTCGTCCCTCGCGCGCTACCGCAGACATTCAGGTG	2400
Db	2341	CGGTCCCGGCTTCTCAAGCAAACTCGTCCCTCGCGCGCTACCGCAGACATTCAGGTG	2400
Qy	2401	GCCCATCTACAGCGCCCTACTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA	2460
Db	2401	GCCCATCTACAGCGCCCTACTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA	2460
Qy	2461	GCCCAAGGTTATAGGTGCTTCTGCTGAAACCGCTCGCGCGCACTTACGCTGCGG	2520
Db	2461	GCCCAAGGTTATAGGTGCTTCTGCTGAAACCGCTCGCGCGCACTTACGCTGCGG	2520
Qy	2521	GCCTATATGTTAAGGCACATGTTATCGACCTAAACATCAGAACCGGGTAAAGCAATC	2580
Db	2521	GCCTATATGTTAAGGCACATGTTATCGACCTAAACATCAGAACCGGGTAAAGCAATC	2580
Qy	2581	ACACCGGTCGCCCATCAGTCTCTGCACTCTGCACTATGGCAAGTTCTTTCGCGCGGTTG	2640
Db	2581	ACACCGGTCGCCCATCAGTCTCTGCACTCTGCACTATGGCAAGTTCTTTCGCGCGGTTG	2640
Qy	2641	TCTGGGCGGCTTATGATCATTAATATGATGAGTGCACCTCACTCACTGATGACCACT	2700
Db	2641	TCTGGGCGGCTTATGATCATTAATATGATGAGTGCACCTCACTCACTGATGACCACT	2700
Qy	2701	ATCTGGGCACTCGGCACTCTCTGCACTCTGCACTATGGCAAGTTCTTTCGCGCGGTTG	2760
Db	2701	ATCTGGGCACTCGGCACTCTCTGCACTCTGCACTATGGCAAGTTCTTTCGCGCGGTTG	2760
Qy	2761	CTGCCACCGCTACGCTCGGGATCGGTGACCGGCTGAGCGGCTGAGCGGCTGCTGCTG	2820
Db	2761	CTGCCACCGCTACGCTCGGGATCGGTGACCGGCTGAGCGGCTGAGCGGCTGCTGCTG	2820
Qy	2821	GCTCTGTCAGCACTGAGAAATCCCTTTTATGGAAGGCAATCCCATTCGAGAGGAGT	2880
Db	2821	GCTCTGTCAGCACTGAGAAATCCCTTTTATGGAAGGCAATCCCATTCGAGAGGAGT	2880
Qy	2881	AAGGGGGGAGGACCTCTCTTCTGCGCTTCCCAAGAAATGATGATGATGATGATGATG	2940
Db	2881	AAGGGGGGAGGACCTCTCTTCTGCGCTTCCCAAGAAATGATGATGATGATGATGATG	2940



D	b	5101	GATGGGTCGGGTGCACAGGTACGCTCCAGCGTSCAAACCCTCTCTACGGGAGGAGTC	5160
Q	y	5161	ACATTCTCGTTCGGCTCAATCAATACCTGTTGGTTCACAGCTCCCATCGAGCCCGAA	5220
D	b	5161	ACATTCTCGTTCGGCTCAAACAATCCTGGTGGGTACAGCTCCCATCGAGCCCGAA	5220
Q	y	5221	CGGACGTAGCAGTGCTCACTTCCANGTCTCACGACCCCTCCCACATTAACGGCGAGACG	5280
D	b	5221	CGGACGTAGCAGTGCTCACTTCCATGTCTACCGACCCCTCCCACATTAACGGCGAGACG	5280
Q	y	5281	GCTAAGCGTAGCGTCGGCCAGGGGATCTCCGCCCTCCTTGGCCAGCTCATCAGCTAGCGAG	5340
D	b	5281	GCTAAGCGTAGCGTCGGCCAGGGGATCTCCCCCTCTCTTGGCCAGCTCATCAGCTAGCGAG	5340
Q	y	5341	CTGTCTGCCCTTCCTTGAAGGCAACATGCATACTCCCGTCACTACTCCTCCGAGCGCTGAC	5400
D	b	5341	CTGTCTGCCCTTCCTTGAAGGCAACATGCATACTACCGCTCATGACTCCCGAGCGCTGAC	5400
Q	y	5401	CTCATCGAGGCCAACCTCCTGTGGCGGAGGAGATGGCGGGGAATCAACCCCGCTGGAG	5460
D	b	5401	CTCATCGAGGCCAACCTCCTGTGGCGGAGGAGATGGCGGGGAATCAACCCCGCTGGAG	5460
Q	y	5461	TCAGAAAATAAGGTAGTAATTTTGAGACTCTTTCAGCGCGCTCCAAGCGGAGGAGTAG	5520
D	b	5461	TCAGAAAATAAGGTAGTAATTTTGAGACTCTTTCAGCGCGCTCCAAGCGGAGGAGTAG	5520
Q	y	5521	AGGGAAGTATCCGTTCCGGCGGAGATCCTGCGAGGTCACAGGAATTCCTCCAGCCGATG	5580
D	b	5521	AGGGAAGTATCCGTTCCGGCGGAGATCCTGCGAGGTCACAGGAATTCCTCCAGCCGATG	5580
Q	y	5581	CCCATATGGGACACGCCGGATTACAACTTCACTGTTAGAGTCTCTGGAAGGACCCGGAC	5640
D	b	5581	CCCATATGGGACACGCCGGATTACAACTTCACTGTTAGAGTCTCTGGAAGGACCCGGAC	5640
Q	y	5641	TACGTCCTCCAGTGATACCGGCTGCCATTGCGCGCTGCCAAGGCGCCCTCCGATACCA	5700
D	b	5641	TACGTCCTCCAGTGATACCGGCTGCCATTGCGCGCTGCCAAGGCGCCCTCCGATACCA	5700
Q	y	5701	CCTCCACGGAGGAAGAGCGGTTGTCTGTGAGAATCTACCGTGCTTCTGCTCTGGCG	5760
D	b	5701	CCTCCACGGAGGAAGAGCGGTTGTCTGTGAGAATCTACCGTGCTTCTGCTCTGGCG	5760
Q	y	5761	GAGCTCGGCACAAAGACTCTCGGAGCTCCGAATCTGTCGCGCGTTCGACACGGCACGGCA	5820
D	b	5761	GAGCTCGGCACAAAGACTCTCGGAGCTCCGAATCTGTCGCGCGTTCGACACGGCACGGCA	5820
Q	y	5821	ACGGCTCTCTGTACACAGCCCTCCGACGCGGCGAGCGGGATCCGACGTTGATCGTAC	5880
D	b	5821	ACGGCTCTCTGTACACAGCCCTCCGACGCGGCGAGCGGGATCCGACGTTGATCGTAC	5880
Q	y	5881	TCCTCCATGCCCTTGAAGGGAGCCCGGGGATCCCGATCTCAGCGACGGGCTCTGG	5940
D	b	5881	TCCTCCATGCCCTTGAAGGGAGCCCGGGGATCCCGATCTCAGCGACGGGCTCTGG	5940
Q	y	5941	TCTACCGTAGCAGGAGGCTAGTAGGACGTGCTGTGCTGTGCTGTGATCTCACATGG	6000
D	b	5941	TCTACCGTAGCAGGAGGCTAGTAGGACGTGCTGTGCTGTGCTGTGATCTCACATGG	6000
Q	y	6001	ACAGGCGCCCTGATACGCCATCGCTCGGAGGAACCAAGCTGCCCATCAATGCACTG	6060
D	b	6001	ACAGGCGCCCTGATACGCCATCGCTCGGAGGAACCAAGCTGCCCATCAATGCACTG	6060
Q	y	6061	AGCAACTCTTTGCTCCGTACCCAACTTGTCTATGCTACAACTCTCGACGCGCAAGC	6120
D	b	6061	AGCAACTCTTTGCTCCGTACCCAACTTGTCTATGCTACAACTCTCGACGCGCAAGC	6120
Q	y	6121	CTGGGACAGAAGGTCACCTTTGACAGACTCGAGCTCTGACGACCACTACCGGGAC	6180
D	b	6121	CTGGGACAGAAGGTCACCTTTTGAAGACTCGAGCTCTGACGACCACTACCGGGAC	6180
Q	y	6181	GTGCTCAAGGAGATGAGGCGAGGCGTCCACAGTTAAGGCTAACTTATCCGTGGAG	6240
D	b	6181	GTGCTCAAGGAGATGAGGCGAGGCGTCCACAGTTAAGGCTAACTTATCCGTGGAG	6240



Qy	1021	TGGGTGGCGGACCGCTATTACAGACATAGCGTTGGCTACCGTGATATTGCTCAAGAGC	1081
Db	1021	TGGGTGGCGGACCGCTATTACAGACATAGCGTTGGCTACCGTGATATTGCTCAAGAGC	1080
Qy	1081	TTGCGCGCGAATGGGCTGACCGCTTCCTCGTGTCTTACGGTATCGCGCTCCCGATTTCG	1140
Db	1081	TTGCGCGCGAATGGGCTGACCGCTTCCTCGTGTCTTACGGTATCGCGCTCCCGATTTCG	1140
Qy	1141	AGCGCATCGCTTCTATTCGCTTCTTGACGAGTTCTTCTGAGTTTAAACAGACACAACG	1200
Db	1141	AGCGCATCGCTTCTATTCGCTTCTTGACGAGTTCTTCTGAGTTTAAACAGACACAACG	1200
Qy	1201	GTTTCCCTTAGCGGGATCAATTCCGCCCTCTCCCTCCCGCCCTACGTTACTGGC	1260
Db	1201	GTTTCCCTTAGCGGGATCAATTCCGCCCTCTCCCTCCCGCCCTACGTTACTGGC	1260
Qy	1261	CGAAGCGCTTGGAAATGAAGCCGGTGTGCGTTGTCTATATGTATTATTTCCACCATTTG	1320
Db	1261	CGAAGCGCTTGGAAATGAAGCCGGTGTGCGTTGTCTATATGTATTATTTCCACCATTTG	1320
Qy	1321	CCGTCCTTTGGCAATGTGAGGGCCGGAACCTGGCCCTGTCCTTCTTGAAGCATTCCT	1380
Db	1321	CCGTCCTTTGGCAATGTGAGGGCCGGAACCTGGCCCTGTCCTTCTTGAAGCATTCCT	1380
Qy	1381	AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGCTGTGTTGAATGCGTGAAGGAAGCA	1440
Db	1381	AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGCTGTGTTGAATGCGTGAAGGAAGCA	1440
Qy	1441	GTTCTCTGGAAGTCTTTGAAGACAAACAGCTCTGTAGCGACCTTTTCGAGCGACGG	1500
Db	1441	GTTCTCTGGAAGTCTTTGAAGACAAACAGCTCTGTAGCGACCTTTTCGAGCGACGG	1500
Qy	1501	AACCCCCACCTGCGCACAGTGCCTCTGCGGCCAAGACCAAGTGTATGAAGTACACCT	1560
Db	1501	AACCCCCACCTGCGCACAGTGCCTCTGCGGCCAAGACCAAGTGTATGAAGTACACCT	1560
Qy	1561	GCAAAGCGGCAACAACCCAGTCGACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAA	1620
Db	1561	GCAAAGCGGCAACAACCCAGTGCACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGGTATTTCAACAGGGGCTGAAGGATCCCAAGAGTACCCCATTTGT	1680
Db	1621	TGGCTCTCTCAAGCGTATTTCAACAGGGGCTGAAGGATCCCAAGAGTACCCCATTTGT	1680
Qy	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGCTTTAGTTCGAGTTAAAA	1740
Db	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGCTTTAGTTCGAGTTAAAA	1740
Qy	1741	AAGCTCTAGGCCCCCGGAACAACGGGACGCTGTTTTCCTTCAAAAAACAGATAATACC	1800
Db	1741	AAGCTCTAGGCCCCCGGAATCAACGGGACGCTGTTTTCCTTGAANAACAGATAATACC	1800
Qy	1801	ATGGCGCTATTAGGGCTACTCCGAACAGCGAGGCTACTTGGTGCAATCATCT	1860
Db	1801	ATGGCGCTATTAGGGCTACTCCCAACACAGCGAGGCTACTTGGTGCAATCATCT	1860
Qy	1861	AGGCTCACAGCGGGACAGGAACAGGTGCGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Db	1861	AGGCTCACAGCCGGGACAGGAACAGGTGCGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Qy	1921	ACACAATCTTCTGCGACCTCGCTCAATGGCGTGTGTGACTGCTCTATCATGTGTGC	1980
Db	1921	ACACAATCTTCTGCGACCTCGGTCAATGGCGTGTGTGACTGCTCTATCATGTGTGC	1980
Qy	1981	GGTCTCAAGACCTTGGCGGCCCAAGGGCCCAATCAACCAATGTACCAATGTGGAC	2040
Db	1981	GGTCTCAAGACCTTGGCGGCCCAAGGGGCCCAATCAACCAATGTACCAATGTGGAC	2040
Qy	2041	CAGGACTCGTGGCTGGCAAGCGCCCGGGGGCGTTCCTTGACACATGCACCTGC	2100
Db	2041	CAGGACTCGTGGCTGGCAAGCGCCCGGGGGCGTTCCTTGACACATGCACCTGC	2100

QY	2101	GGCAGCTCGGACCTTTACTGGTACAGAGGATCCGATGTCATTCGGTGGCGCGCGGG	2160
DB	2101	GGCAGCTCGGACCTTTACTGGTACAGAGGCATCCGATGTGTCATTCGGTGGCGCGCGGG	2160
QY	2161	GGCGACAGCAGGGGAGCCTACTCTCCCCAGGCCCGCTCTCCTACTTTGAAGGGCTCTTCG	2220
DB	2161	GGCGACAGCAGGGGAGCCTACTCTCCCCAGGCCCGCTCTCCTACTTTGAAGGGCTCTTCG	2220
QY	2221	GGCGGTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCATCTTTCGGGCTCGCGTGTGC	2280
DB	2221	GGCGGTCCACTGCTCTGCCCTCGGGGCAAGCTGTGGGCATCTTTCGGGCTCGCGTGTGC	2280
QY	2281	ACCCGAGGGTTGCGAAGCGGTGGACTTTTGTACCCGTCGAGTCTATGCGAAACCACTATG	2340
DB	2281	ACCCGAGGGTTGCGAAGCGGTGGACTTTTGTACCCGTCGAGTCTATGCGAAACCACTATG	2340
QY	2341	CGGTCCCGGCTTTCACGGACAACCTCGTCCCTCCGSCCGTACCGGAGACATTTCCAGGTG	2400
DB	2341	CGGTCCCGGCTTTCACGGACAACCTCGTCCCTCCGSCCGTACCGGAGACATTTCCAGGTG	2400
QY	2401	GCCCATCTACAGCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCCGGCTGGGTATGCA	2460
DB	2401	GCCCATCTACAGCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCCGGCTGGGTATGCA	2460
QY	2461	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGTCGTCGCGGCACCCCTAGGTTTCGGG	2520
DB	2461	GCCCAAGGTTAAGGTGCTTGTCTGAACCCGTCGTCGCGGCACCCCTAGGTTTCGGG	2520
QY	2521	GCGTATATGCTTAAGGCACATGTGTACGCCCTAAATCAGAACCGGGTAAAGCAATC	2580
DB	2521	GCGTATATGCTTAAGGCACATGTGTATCGCCCTAAATCAGAACCGGGTAAAGCAATC	2580
QY	2581	ACCAACGGTGCCTCCATCAGCTACTCCACCTATGGCAAGTTTCTGCGACGGTGGTTCG	2640
DB	2581	ACCAACGGTGCCTCCATCAGCTACTCCACCTATGGCAAGTTTCTGCGACGGTGGTTCG	2640
QY	2641	TCGCGGGCGCCTATGACATCATATATGTAGTGAGTGCCACTCAACTGACTCGACCACT	2700
DB	2641	TCGCGGGCGCCTATGACATCATATATGTAGTGAGTGCCACTCAACTGACTCGACCACT	2700
QY	2701	ATCCTGGGCATCGGCACAGTCTGGACCAAGCGGAGACGGCTGGAGGGGACATCTGCTGGT	2760
DB	2701	ATCCTGGGCATCGGCACAGTCTGGACCAAGCGGAGACGGCTGGAGGGGACATCTGCTGGT	2760
QY	2761	CTCGCCACCGCTACGGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
DB	2761	CTCGCCACCGCTACGGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
QY	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTATGGCAAGGCCATCCCATCGAGACCATC	2880
DB	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTATGGCAAGGCCATCCCATCGAGACCATC	2880
QY	2881	AAGGGGGGAGGCACCTCATTTTCTGCCATTCGAAGAAGAAATGTGATGAGCTCGCGCG	2940
DB	2881	AAGGGGGGAGGCACCTCATTTTCTGCCATTCGAAGAAGAAATGTGATGAGCTCGCGCG	2940
QY	2941	AAGCTGTCCGGCCTCGGACTCAATGTGTAGCATATTACCGGGGCTTCATGTATCCGTC	3000
DB	2941	AAGCTGTCCGGCCTCGGACTCAATGTGTAGCATATTACCGGGGCTTCATGTATCCGTC	3000
QY	3001	ATACCAACTAGCGGAGCGTCAATGTTCGTAGCAACCGACGCTCTAATGACGGGCTTTAC	3060
DB	3001	ATACCAACTAGCGGAGCGTCAATGTTCGTAGCAACCGACGCTCTAATGACGGGCTTTAC	3060
QY	3061	GGCGATTTGCACTCACTGATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAGC	3120
DB	3061	GGCGATTTGCACTCACTGATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAGC	3120
QY	3121	CTGGAACCGACCTTACCATTGAGACGACGCGGTGCCAAGACGCGGTGTACGCTCG	3180
DB	3121	CTGGAACCGACCTTACCATTGAGACGACGCGGTGCCAAGACGCGGTGTACGCTCG	3180
QY	3181	CAGCGCGAGGACGACTCGTATGGGCGAGGATGGGCATTTACAGGTTTGTGACTCCAGGA	3240

Db 3181 CAGCGCGAGGAGGAGCTGGTAGGGCAGGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
 QY 3241 GAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGGAGTGTATACGCGGGCTGT 3300  
 Db 3241 GAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGGAGTGTATACGCGGGCTGT 3300  
 QY 3301 GCTTGTGTAGAGCTCAGCGCCGCGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
 Db 3301 GCTTGTGTAGAGCTCAGCGCCGCGAGACCTCAGTTAGTTTGGGGCTTACCTAAACACA 3360  
 QY 3361 CAGAGGTTCCCGTCTGCGAGACCATCTGGAGTTCTGGAGAGCGCTCTTACAGGCCTC 3420  
 Db 3361 CAGAGGTTCCCGTCTGCGAGACCATCTGGAGTTCTGGAGAGCGCTCTTACAGGCCTC 3420  
 QY 3421 ACCCATAGAGCCCAATTTCTGTCCAGACTTAAGCAGGAGAGCAACTTCCCTTAC 3480  
 Db 3421 ACCCATAGAGCCCAATTTCTGTCCAGACTTAAGCAGGAGAGCAACTTCCCTTAC 3480  
 QY 3481 CTGGTAGCATACAGGCTAGGTTGCGCAGGCTCAGGCTCAGCTCCACCTCATCGTGGGAC 3540  
 Db 3481 CTGGTAGCATACAGGCTACGTTGCGCAGGCTCAGGCTCAGCTCCACCTCATCGTGGGAC 3540  
 QY 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCTTACGCTGCAAGGCTCAGGCTCCTGCTG 3600  
 Db 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCTTACGCTGCAAGGCTCAGGCTCCTGCTG 3600  
 QY 3601 TATAGGCTGGAGCCGTTCAAACAGAGTTTACCAACACACCCCATTAACCAATACATC 3660  
 Db 3601 TATAGGCTGGAGCCGTTCAAACAGAGTTTACCAACACACCCCATTAACCAATACATC 3660  
 QY 3661 ATGGCATGATGTGCGCTGACCTGAGGTCGTCAAGCAGCAGCTGGTGTGTTAGCGGA 3720  
 Db 3661 ATGGCATGATGTGCGCTGACCTGAGGTCGTCAAGCAGCAGCTGGTGTGTTAGCGGA 3720  
 QY 3721 GTCTTAGCAGCTCTGCGCGGCTATTCCTGACAAAGCAGCGCTGTGTCATTGTGGGCAGG 3780  
 Db 3721 GTCTTAGCAGCTCTGCGCGGCTATTCCTGACAAAGCAGCGCTGTGTCATTGTGGGCAGG 3780  
 QY 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGGAGTTT 3840  
 Db 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGGAGTTT 3840  
 QY 3841 GATGAGATGGAAGTGGCTCAGACCTCCCTTACATCGAAAGGGAATGCGAGCTGCC 3900  
 Db 3841 GATGAGATGGAAGTGGCTCAGACCTCCCTTACATCGAAAGGGAATGCGAGCTGCC 3900  
 QY 3901 GAACAAATTCAGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGGAGTTT 3960  
 Db 3901 GAACAAATTCAGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTTACCGGGAGTTT 3960  
 QY 3961 GCTGCTCCGTTGTGGAATCCAGTGGCGAACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
 Db 3961 GCTGCTCCGTTGTGGAATCCAGTGGCGAACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
 QY 4021 TGAATTTATCAGCGGATACATATTTAGAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
 Db 4021 TGAATTTATCAGCGGATACATATTTAGAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
 QY 4081 GCGATAGCATCACTGATGGCAATTCAGAGCTTATCAACAGCCCGCTCAGCACCACAT 4140  
 Db 4081 GCGATAGCATCACTGATGGCAATTCAGAGCTTATCAACAGCCCGCTCAGCACCACAT 4140  
 QY 4141 ACCCTCTGTTTAACTCTTGGGGGATGGTGGCGGCCCACTTGTCTCTCCAGGCT 4200  
 Db 4141 ACCCTCTGTTTAACTCTTGGGGGATGGTGGCGGCCCACTTGTCTCTCCAGGCT 4200  
 QY 4201 GCTTCTGCTTTTGTAGCGCGGCTCAGCTGGAGCGGCTGTGGAGCATAGGCTTTGGG 4260  
 Db 4201 GCTTCTGCTTTTGTAGCGCGGCTCAGCTGGAGCGGCTGTGGAGCATAGGCTTTGGG 4260  
 QY 4261 AAGGTGCTTGTGGAATTTTGGCAGGTTATGGAGAGGGGTGGCAGCGGCTCGTGCC 4320

Db 4261 AAGGTGCTTGTGGAATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGGCTCGTGCC 4320  
 QY 4321 TTTAAGGTCATCAGCGGCGAGATGCCCTCCACCGAGGACCTGTTAACTTCCCTGCT 4380  
 Db 4321 TTTAAGGTCATCAGCGGCGAGATGCCCTCCACCGAGGACCTGTTAACTTCCCTGCT 4380  
 QY 4381 ATCTCTCTCCCTCGGGCCCTAGTCTCGGGGTCTGTGCGCAGCGATCTTGGGTGCGCAC 4440  
 Db 4381 ATCTCTCTCCCTCGGGCCCTAGTCTCGGGGTCTGTGCGCAGCGATCTTGGGTGCGCAC 4440  
 QY 4441 GTGGGCCAGGAGGAGGGGCTCTGTCAGTGGATGAACCGGCTGANTAGCGTTTGGTTCCGG 4500  
 Db 4441 GTGGGCCAGGAGGAGGGGCTCTGTCAGTGGATGAACCGGCTGANTAGCGTTTGGTTCCGG 4500  
 QY 4501 GGTAAACAGCTCTCCCGCAGCCTATGTCCTGAGAGCAGCGCTCAGCAGCTGTGCACT 4560  
 Db 4501 GGTAAACAGCTCTCCCGCAGCCTATGTCCTGAGAGCAGCGCTCAGCAGCTGTGCACT 4560  
 QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
 Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
 QY 4621 GAGGACTGTCTCAGCGCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
 Db 4621 GAGGACTGTCTCAGCGCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
 QY 4681 ACGGTTTACATGTTTCAAGACCTGCTCAGTCCAGCTCCTGCGGATTTGCCGGA 4740  
 Db 4681 ACGGTTTACATGTTTCAAGACCTGCTCAGTCCAGCTCCTGCGGATTTGCCGGA 4740  
 QY 4741 GTCCCTCTCTCTCATGTCATCAAGGAGTCTGCGGGGCGACCGCATCATG 4800  
 Db 4741 GTCCCTCTCTCTCATGTCATCAAGGAGTCTGCGGGGCGACCGCATCATG 4800  
 QY 4801 CAAACACCTGCGCATGTCGAGCAGATCACCGCATGTGAAACCGGTTCCATGAGG 4860  
 Db 4801 CAAACACCTGCGCATGTCGAGCAGATCACCGCATGTGAAACCGGTTCCATGAGG 4860  
 QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACGTGCAATTCCTTCTAGGCGCTGTGGCGGTG 4920  
 Db 4861 ATCGTGGGCTTAGGACCTGTAGTAAACGTGCAATTCCTTCTAGGCGCTGTGGCGGTG 4920  
 QY 4921 ACCACGGGCTTGTGACGCGCTCCCGGGCGCAATTTCTTCTAGGCGCTGTGGCGGTG 4980  
 Db 4921 ACCACGGGCTTGTGACGCGCTCCCGGGCGCAATTTCTTCTAGGCGCTGTGGCGGTG 4980  
 QY 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGAATTTCCACTACGTGACGGGCGATG 5040  
 Db 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGAATTTCCACTACGTGACGGGCGATG 5040  
 QY 5041 ACCACTGACAACTAAAGTCCCGGCTGTCAGGTTCCCGGCGCAATTTCTTCTACAGAGTG 5100  
 Db 5041 ACCACTGACAACTAAAGTCCCGGCTGTCAGGTTCCCGGCGCAATTTCTTCTACAGAGTG 5100  
 QY 5101 GATGGGCTGCGGTGTCAGAGTACGCTCCAGCGTGCAGAACCCCTCTTACGGGAGGTC 5160  
 Db 5101 GATGGGCTGCGGTGTCAGAGTACGCTCCAGCGTGCAGAACCCCTCTTACGGGAGGTC 5160  
 QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCATCCAGCGCGGAA 5220  
 Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCATCCAGCGCGGAA 5220  
 QY 5221 CCGGAGTACAGTGTCTTCCATGCTCAGCGACCCCTCCACATTTACGGCGGAGACG 5280  
 Db 5221 CCGGAGTACAGTGTCTTCCATGCTCAGCGACCCCTCCACATTTACGGCGGAGACG 5280  
 QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCGAGCTCATCAGCTAGGCG 5340  
 Db 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCGAGCTCATCAGCTAGGCG 5340  
 QY 5341 CTGTCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGCTCCCGGAGCGTGAC 5400  
 Db 5341 CTGTCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGCTCCCGGAGCGTGAC 5400



QY 5401 CTCATCGAGGCAACCTCTCTGTGGCGCAGAGATGGCGGGAAACATACCCCGGTGGAG 5460  
DB 5401 CTCATCGAGGCAACCTCTCTGTGGCGCAGAGATGGCGGGAAACATACCCCGGTGGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
DB 5461 TCAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGGCGAGATCTCTGGGAGGTCCAGAAATTCCTTCGAGCGGATG 5580  
DB 5521 AGGGAAGTATCCGTTCCGGCGAGATCTCTGGGAGGTCCAGAAATTCCTTCGAGCGGATG 5580  
QY 5581 CCCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTCTGGAAGAAATTCCTTCGAGCGGATG 5640  
DB 5581 CCCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTCTGGAAGAAATTCCTTCGAGCGGATG 5640  
QY 5641 TAGCTCCCTCCAGTGGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
DB 5641 TAGCTCCCTCCAGTGGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
QY 5701 CCTTCCAGGAGGAGAGAGCGGTTGTCTGTCTGAGAAATCTACCGTGTCTTCTGCGTGGCG 5760  
DB 5701 CCTTCCAGGAGGAGAGAGCGGTTGTCTGTCTGAGAAATCTACCGTGTCTTCTGCGTGGCG 5760  
QY 5761 GAGCTCGCCCAAAAGACCTTCGGCAGCTCCGAATCGTCCGCGCTCGACAGCGGCACCGCA 5820  
DB 5761 GAGCTCGCCCAAAAGACCTTCGGCAGCTCCGAATCGTCCGCGCTCGACAGCGGCACCGCA 5820  
QY 5821 ACGGCTCTCTGTACACGCTTCGACGAGCGGCGGATCCGACGTTGAGTCTGTAC 5880  
DB 5821 ACGGCTCTCTGTACACGCTTCGACGAGCGGCGGATCCGACGTTGAGTCTGTAC 5880  
QY 5881 TCCTTCATGCCCTTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG 5940  
DB 5881 TCCTTCATGCCCTTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG 5940  
QY 5941 TCTACCGTAAAGGAGGAGTGTAGGAGCGTCTGTCTGCTGCTCGATGCTCTACAGATG 6000  
DB 5941 TCTACCGTAAAGGAGGAGTGTAGGAGCGTCTGTCTGCTGCTCGATGCTCTACAGATG 6000  
QY 6001 ACAGGCGCTCTGATCAGCCATCGCTTCGAGGAGAAACCAAGCTGCCATCAATGCACTG 6060  
DB 6001 ACAGGCGCTCTGATCAGCCATCGCTTCGAGGAGAAACCAAGCTGCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCGTCCGTCACCAACTTGGTCTATGCTACCAATCTCGAGCGCAAGC 6120  
DB 6061 AGCAACTCTTTGCTCGTCCGTCACCAACTTGGTCTATGCTACCAATCTCGAGCGCAAGC 6120  
QY 6121 CTGCGGAGAGAGGTACCTTTGACAGACTCGAGTCTCTGGACGACCACTACCGGAC 6180  
DB 6121 CTGCGGAGAGAGGTACCTTTGACAGACTCGAGTCTCTGGACGACCACTACCGGAC 6180  
QY 6181 GTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAAGCTAACTCTATCCGTGGAG 6240  
DB 6181 GTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAAGCTAACTCTATCCGTGGAG 6240  
QY 6241 GAAGCTGTAAAGTACGCGCCCAATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
DB 6241 GAAGCTGTAAAGTACGCGCCCAATTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
QY 6301 GAGTCCGGAACCTATCCAGCAGGCGTTAAACCATTCGCTCCGTTGCGAAGACTTG 6360  
DB 6301 GAGTCCGGAACCTATCCAGCAGGCGTTAAACCATTCGCTCCGTTGCGAAGACTTG 6360  
QY 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGCAAAATGAGGTTTCTGC 6420  
DB 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGCAAAATGAGGTTTCTGC 6420  
QY 6421 GTCCAAACGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTCAGATTTGGGG 6480  
DB 6421 GTCCAAACGAGAGGGGGCGCAAGCCAGCTCGCTTATCGTATTCAGATTTGGGG 6480

QY 6481 GTTCGTGTGTGCGAGAAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG 6540  
DB 6481 GTTCGTGTGTGCGAGAAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG 6540  
QY 6541 ATGGGCTCTTCATACGGATTCCAATCTCTCTGACAGCGGTCGAGTCTCTGGTGAAT 6600  
DB 6541 ATGGGCTCTTCATACGGATTCCAATCTCTCTGACAGCGGTCGAGTCTCTGGTGAAT 6600  
QY 6601 GCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGATATGACACCCGCTGTGTTGACTCA 6660  
DB 6601 GCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGATATGACACCCGCTGTGTTGACTCA 6660  
QY 6661 ACGGTCTACTGAGATGATCATCCGTTGTTGAGGAGTCAATCTACCAATGTTGTTGACTG 6720  
DB 6661 ACGGTCTACTGAGATGATCATCCGTTGTTGAGGAGTCAATCTACCAATGTTGTTGACTG 6720  
QY 6721 CCCGAAGCAGACAGGCGCATAAAGTCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 6780  
DB 6721 CCCGAAGCAGACAGGCGCATAAAGTCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 6780  
QY 6781 ACTAATTTCTAAAGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGTACTGACG 6840  
DB 6781 ACTAATTTCTAAAGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGTACTGACG 6840  
QY 6841 ACCAGCTGCGGTAATACCTTCAATGTTACTTGAAGCGCGCTGCGGCTGTCTGAGCTGCG 6900  
DB 6841 ACCAGCTGCGGTAATACCTTCAATGTTACTTGAAGCGCGCTGCGGCTGTCTGAGCTGCG 6900  
QY 6901 AAGCTTCAGGACTGACGATGCTCGTATGCGGAGACCACTTGTGCTTATCTGTGAAAGC 6960  
DB 6901 AAGCTTCAGGACTGACGATGCTCGTATGCGGAGACCACTTGTGCTTATCTGTGAAAGC 6960  
QY 6961 GCGGGGACCCAAAGAGGAGCGGAGCTTACGGGCTTTCAGCGAGGCTATGACTAGATAC 7020  
DB 6961 GCGGGGACCCAAAGAGGAGCGGAGCTTACGGGCTTTCAGCGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTCTGGGAGCCGCCAAACAGATATACGACTTGGAGTTGATTAACATATGC 7080  
DB 7021 TCTGCCCCCTCTGGGAGCCGCCAAACAGATATACGACTTGGAGTTGATTAACATATGC 7080  
QY 7081 TCCTTCAATGTGTTCAGTCGCGCATCTGGGCAAAAGGTTGATCTATCTCACCCGT 7140  
DB 7081 TCCTTCAATGTGTTCAGTCGCGCATCTGGGCAAAAGGTTGATCTATCTCACCCGT 7140  
QY 7141 GACCCCAACACCCCTTTCGCGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
DB 7141 GACCCCAACACCCCTTTCGCGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
QY 7201 TCCTGGCTAGGCAACATCATCTATGCGCCCACTTGTGGGCAAGGATGATCTGATG 7260  
DB 7201 TCCTGGCTAGGCAACATCATCTATGCGCCCACTTGTGGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTTCTCCATCTCTTAGCTCAGGAAACAACTTGAAAGCCCTAGATTTGTCAG 7320  
DB 7261 ACTCATTTCTTCTCCATCTCTTAGCTCAGGAAACAACTTGAAAGCCCTAGATTTGTCAG 7320  
QY 7321 ATCTACGGGCTGTGTTACTTCAATTGAGCCACTTGAACCTCAGATCAATCAACGACTC 7380  
DB 7321 ATCTACGGGCTGTGTTACTTCAATTGAGCCACTTGAACCTCAGATCAATCAACGACTC 7380  
QY 7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
DB 7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
QY 7441 TCATGCTCAGGAAACTTGGGGTACCGGCTTGGAGTCTGGAGATCGGGCCAGAGT 7500  
DB 7441 TCATGCTCAGGAAACTTGGGGTACCGGCTTGGAGTCTGGAGATCGGGCCAGAGT 7500  
QY 7501 GTCCGCGCTAGGCTACTGTCCCAAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
DB 7501 GTCCGCGCTAGGCTACTGTCCCAAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGCAAGTAAAGCAAGCTCAAACTCACTCCAAATCCCGCTCGCTGCCAGTTGGAT 7620



Db	7561	AACTGGGCAGTAAAGCAAGCTCAAACTCACTCCAAATCCCGGCTGCTCCAGTTGGAT	7620
Qy	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGAGACATATACACGCTGTCTCGT	7680
Db	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGAGACATATACACGCTGTCTCGT	7680
Qy	7681	GCCCGACCCCGCTGGTTTCATGTGGTGCTACTCTCTACTTTCTGTAGGGGTAGGCATCTAT	7740
Db	7681	GCCCGACCCCGCTGGTTTCATGTGGTGCTACTCTCTACTTTCTGTAGGGGTAGGCATCTAT	7740
Qy	7741	CTACTCCCCAACCGATGAACGGGAGACCTAAACACTCCAGGCCAATAGGCCATCCTGTTTT	7800
Db	7741	CTACTCCCCAACCGATGAACGGGAGACTAAACACTCCAGGCCAATAGGCCATCCTGTTTT	7800
Qy	7801	TTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCCTTT	7860
Db	7801	TTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCCTTT	7860
Qy	7861	TTTTTCTCTTTTTTTTTTCTCTTTCTCTTCTCTTTGGTGGCTCAATCTTAGCCCTAGTCACGGC	7920
Db	7861	TTTTTCTCTTTTTTTTTTCTCTTTCTCTTCTCTTTGGTGGCTCAATCTTAGCCCTAGTCACGGC	7920
Qy	7921	TAGCTGTCAAAAGTCCGTGAGCCGCTTGACTTCGACAGAGTGCTGATACTGGCCCTCTCTGC	7980
Db	7921	TAGCTGTCAAAAGTCCGTGAGCCGCTTGACTTCGACAGAGTGCTGATACTGGCCCTCTCTGC	7980
Qy	7981	AGATCAAGT	7989
Db	7981	AGATCAAGT	7989

```

RESULT 13
US-10-434-842-17
; Sequence 17, Application US/10434842
; Publication No. US20040005549A1
; GENERAL INFORMATION:
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION
; FILE REFERENCE: 0342/1H395US3
; CURRENT APPLICATION NUMBER: US/10/434,842
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/233,307
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 10/005,469
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: FCA22 Replicon Sequence
US-10-434-842-17

```

	Query Match	Best Local Similarity	99.9%; Matches 7985;	Score 7982.6;	DB 16;	Length 7992;				
		99.9%; Matches 7985;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
QY	1	GCACGCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTCTGAGGA	CTCTG	60						
DB	1	GCACGCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTCTGAGGA	CTCTG	60						
QY	61	TCCTTACGCAGAAACGGCTTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCCTCCAGGAC	120							
DB	61	TCCTTACGCAGAAACGGCTTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCCTCCAGGAC	120							
QY	121	CCCCCTCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG	180							
DB	121	CCCCCTCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCAG	180							

QY	181	GACGACCGGGTCTCTTTCTTGATCAACCCGCTCAATGCTCGAGATTTGGGCGTGCCTCC	240
Db	181	GACGACCGGGTCTCTTTCTTGATCAACCCGCTCAATGCTCGAGATTTGGGCGTGCCTCC	240
QY	241	GCGAGACTCTAGCGAGTAGTGTGGTGCAGAAAGGCTTGTGGTACTGTGCTGATAGG	300
Db	241	GCGAGACTCTAGCGAGTAGTGTGGTGCAGAAAGGCTTGTGCTACTGTCTGATAGG	300
QY	301	GTCTTGCAGTGCCTCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGATCTCTAAAC	360
Db	301	GTCTTGCAGTGCCTCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGATCTCTAAAC	360
QY	361	CTCAAGAAACCAAGAGCGCGCATGATTGAACAAGATGGATTGACGACGAGTCTC	420
Db	361	CTCAAGAAACCAAGAGCGCGCATGATTGAACAAGATGGATTGACGACGAGTCTC	420
QY	421	CGCGCGCTTGGGTGGAGAGGCTATTCCGGTATGACTGGGCACAAACAGACAATCGGTGCT	480
Db	421	CGCGCGCTTGGGTGGAGAGGCTATTCCGGTATGACTGGGCACAAACAGACAATCGGTGCT	480
QY	481	CTGATCCGCGGTTCCTCGGCTGTGAGCGAGGGCGCCGGTCTTTTGTCAAGACCG	540
Db	481	CTGATCCGCGGTTCCTCGGCTGTGAGCGAGGGCGCCGGTCTTTTGTCAAGACCG	540
QY	541	ACCTGTCCGGTGCCTTGAATGAATCTCAGACGAGGACGCGGCTTATCGTGGCTGCACA	600
Db	541	ACCTGTCCGGTGCCTTGAATGAATCTCAGACGAGGACGCGGCTTATCGTGGCTGCACA	600
QY	601	CGACGGCGTTCCTTCGGCAGCTGTCTCGACGTTGTCTGTGAAGCGGGAAGGACTGGC	660
Db	601	CGACGGCGTTCCTTCGGCAGCTGTCTCGACGTTGTCTGTGAAGCGGGAAGGACTGGC	660
QY	661	TGCTATTGGCGAAGTGC CGGGCAGGATCTCTGTCACTCACTTGTCTCTGCCGAGA	720
Db	661	TGCTATTGGCGAAGTGC CGGGCAGGATCTCTGTCACTCACTTGTCTCTGCCGAGA	720
QY	721	AAGTATCCATCATGGCTGATGCAATCGGGGTGCATACGTTGTATTCGGCTACTGCC	780
Db	721	AAGTATCCATCATGGCTGATGCAATCGGGGTGCATACGTTGTATTCGGCTACTGCC	780
QY	781	CATTTCAGCACCAAGCGAACAATCGCATCGAGCGACAGTACTCGGATGGAGCGGTC	840
Db	781	CATTTCAGCACCAAGCGAACAATCGCATCGAGCGACAGTACTCGGATGGAGCGGTC	840
QY	841	TTGTTCGATCAGGATGATCTGGAAGAAAGAGCATCAGGGGTTCGGCCAGCGGAATCTTCG	900
Db	841	TTGTTCGATCAGGATGATCTGGAAGAAAGAGCATCAGGGGTTCGGCCAGCGGAATCTTCG	900
QY	901	CCAGGCTCAAGGCGCGCATCCCGAGCGAGAGTCTCTGCGTGAACCATGGCGATGCCT	960
Db	901	CCAGGCTCAAGGCGCGCATCCCGAGCGAGAGTCTCTGCTGTAACCATGGCGATGCCT	960
QY	961	GCTTGCAGATATCATGGTGGAAATGGCGCTTTTCTGGATTTCATCGACTGGCGCGC	1020
Db	961	GCTTGCAGATATCATGGTGGAAATGGCGCTTTTCTGGATTTCATCGACTGGCGCGC	1020
QY	1021	TGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATTTGCTGAAGAGC	1080
Db	1021	TGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATTTGCTGAAGAGC	1080
QY	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTTCG	1140
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTTCG	1140
QY	1141	AGCGCATCGCTTCTATTCGCTTCTTGAACAGTTCCTCTGAGTTTAAACAGACCAACG	1200
Db	1141	AGCGCATCGCTTCTATTCGCTTCTTGAACAGTTCCTCTGAGTTTAAACAGACCAACG	1200
QY	1201	GTTTCCCTCTAGCGGATCAATTCGGCCCTCTCTCCCTCCCGCCCTTAAGTTACTGGC	1260
Db	1201	GTTTCCCTCTAGCGGATCAATTCGGCCCTCTCTCCCTCCCGCCCTTAATTTACTGGC	1260

QY 1261 CGAAGCCGCTTGGAAATAAGCGCGGTGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
DB 1261 CGAAGCCGCTTGGAAATAAGCGCGGTGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACAACCTGCGCCCTGTCTTTGACGAGCAATTCCT 1380  
DB 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACAACCTGCGCCCTGTCTTTGACGAGCAATTCCT 1380  
QY 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGGAAGTCTGTTGAATGTCTGTAAGTGTCTGTAAGGAAGCA 1440  
DB 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGGAAGTCTGTTGAATGTCTGTAAGTGTCTGTAAGGAAGCA 1440  
QY 1441 GTTCTCTGGAAGCTTTCTTGAAGCAAAACAAAGCTCTGTAGGAGCCCTTTGAGGCGAGCGG 1500  
DB 1441 GTTCTCTGGAAGCTTTCTTGAAGCAAAACAAAGCTCTGTAGGAGCCCTTTGAGGCGAGCGG 1500  
QY 1501 AACCCCCACCTGGCGCAGGTGCTCTGCGGCCAAAAGCCAAAGCGTGTATAAGATACACCT 1560  
DB 1501 AACCCCCACCTGGCGCAGGTGCTCTGCGGCCAAAAGCGTGTATAAGATACACCT 1560  
QY 1561 GCMAAGGCGGCACAAACCCAGTGCACAGTGTGTGAGTGTGATAGTGTGTAAGAGTCAAA 1620  
DB 1561 GCMAAGGCGGCACAAACCCAGTGCACAGTGTGTGAGTGTGATAGTGTGTAAGAGTCAAA 1620  
QY 1621 TGGTCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAAGGTACCCCATTTG 1680  
DB 1621 TGGTCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTTACATGTGTTAGTCGAGGTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTTACATGTGTTAGTCGAGGTAAAA 1740  
QY 1741 AACGCTTAGGCGCCCGGACCAACCGGAGCGTGTCTTCCCTTTGAAACACGATATACC 1800  
DB 1741 AACGCTTAGGCGCCCGGACCAACCGGAGCGTGTCTTCCCTTTGAAACACGATATACC 1800  
QY 1801 ATGGCGGCTATTACGGGCTACTCCCAACAGAGCGGAGGCTTGGTGTGATCATCACT 1860  
DB 1801 ATGGCGGCTATTACGGGCTACTCCCAACAGAGCGGAGGCTTGGTGTGATCATCACT 1860  
QY 1861 AGCTCAAGCGCGGACAGGAAACAGGTGAGGCGGAGTCCAAAGTGGTCTCCACCGCA 1920  
DB 1861 AGCTCAAGCGCGGACAGGAAACAGGTGAGGCGGAGTCCAAAGTGGTCTCCACCGCA 1920  
QY 1921 ACACAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGAGTGTCTATCATGTGTC 1980  
DB 1921 ACACAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGAGTGTCTATCATGTGTC 1980  
QY 1981 GGCTCAAAGACCTTTGCGGCGCCAAAGGGGCCAATCAACCAAAATGTACACCAATGTGAC 2040  
DB 1981 GGCTCAAAGACCTTTGCGGCGCCAAAGGGGCCAATCAACCAAAATGTACACCAATGTGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCAATGCACCTGC 2100  
DB 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCAATGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTTGTTGACAGGACATGCGGATGCTATTCGCGTGGCGCGG 2160  
DB 2101 GGCAGCTCGGACCTTTACTTTGTTGACAGGACATGCGGATGCTATTCGCGTGGCGCGG 2160  
QY 2161 GGCAGCAGGAGGGAGGCTACTCTCCCGCAGGCGCGTCTCTACTTTGAAGGGCTCTTCG 2220  
DB 2161 GGCAGCAGGAGGGAGGCTACTCTCCCGCAGGCGCGTCTCTACTTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGTCTCACTGCTCTGCGGCAAGCGTGTGGGCACTTTTCGCGGCTGCGGTGTGC 2280  
DB 2221 GCGGTCTCACTGCTCTGCGGCAAGCGTGTGGGCACTTTTCGCGGCTGCGGTGTGC 2280  
QY 2281 ACCCGAGGGGTGCGAAGGCGGTGGAATTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
DB 2281 ACCCGAGGGGTGCGAAGGCGGTGGAATTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
QY 2341 CCGTCCCGGCTTTCAGGCAACCTCGTCCCTCCGCGCGTACCGGAGCAATTCAGGTG 2400

DB 2341 CCGTCCCGGCTTTCAGGCAACCTCGTCCCTCCGCGCGTACCCGAGACATTCAGGCTG 2400  
QY 2401 GCCCATCTACACGCGCCCTACTGGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
DB 2401 GCCCATCTACACGCGCCCTACTGGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
QY 2461 GCCCAAGGGTAAAGGTGCTTGTCTGAAACCGCTCCGTCCGCGCCACCTAGTTTCGGG 2520  
DB 2461 GCCCAAGGGTAAAGGTGCTTGTCTGAAACCGCTCCGTCCGCGCCACCTAGTTTCGGG 2520  
QY 2521 CGGTATATGTCTAAGGCACATGTTATCGACCTTACATCAGAACCGGGTAAAGACCATC 2580  
DB 2521 CGGTATATGTCTAAGGCACATGTTATCGACCTTACATCAGAACCGGGTAAAGACCATC 2580  
QY 2581 ACCACGGGTGCCCCCATCACGTACTCCACCTATGCGAAGTTTCTTGCCGACGGTGTTCG 2640  
DB 2581 ACCACGGGTGCCCCCATCACGTACTCCACCTATGCGAAGTTTCTTGCCGACGGTGTTCG 2640  
QY 2641 TCTGGGGCGGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB 2641 TCTGGGGCGGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTTGCGCATCGGCACAGTCTCGSACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG 2760  
DB 2701 ATCTTGCGCATCGGCACAGTCTCGSACCAAGCGGAGCGGCTGGAGCGGACTCGTCGTG 2760  
QY 2761 CTCGCCACCGCTACGCGCTCCGGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG 2820  
DB 2761 CTCGCCACCGCTACGCGCTCCGGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG 2820  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCCAGACCATC 2880  
DB 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCCAGACCATC 2880  
QY 2881 AAGGGGGGAGGCGACCTCATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCG 2940  
DB 2881 AAGGGGGGAGGCGACCTCATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCG 2940  
QY 2941 AAGCTGTCCGGCTTCGACCTCAATGTGTAGCATATTACCGGGGCTTGTATGATTCGTC 3000  
DB 2941 AAGCTGTCCGGCTTCGACCTCAATGTGTAGCATATTACCGGGGCTTGTATGATTCGTC 3000  
QY 3001 ATACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
DB 3001 ATACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
QY 3061 GGGGATTTGACACTCAGTGATCGACTCAATACATGTGTCAACCAAGACAGTTCGACTCAGC 3120  
DB 3061 GGGGATTTGACACTCAGTGATCGACTCAATACATGTGTCAACCAAGACAGTTCGACTCAGC 3120  
QY 3121 CTGGACCCGACCTTACCAATTGAGACGACACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
DB 3121 CTGGACCCGACCTTACCAATTGAGACGACACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
QY 3181 CAGCGGCGAGGACAGGACTGTGAGGGCAGGATGGGCAATTTACAGGTTTGTACTCCAGGA 3240  
DB 3181 CAGCGGCGAGGACAGGACTGTGAGGGCAGGATGGGCAATTTACAGGTTTGTACTCCAGGA 3240  
QY 3241 GAACGCGCTTCGGGCAATGTTTCGATTCCTCGTTCGTGCGAGTGTCTATGACGCGGCTGT 3300  
DB 3241 GAACGCGCTTCGGGCAATGTTTCGATTCCTCGTTCGTGCGAGTGTCTATGACGCGGCTGT 3300  
QY 3301 GCTTGTGTACGAGCTCACGCGCCGAGAGCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
DB 3301 GCTTGTGTACGAGCTCACGCGCCGAGAGCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCGTCTGCCAGGACCACTCTGGAGTCTTGGAGAGCGTCTTTTACAGGCTC 3420  
DB 3361 CCAGGTTGCCGTCTGCCAGGACCACTCTGGAGTCTTGGAGAGCGTCTTTTACAGGCTC 3420  
QY 3421 ACCCAATAGACGCCCATTTCTTGTCCAGACTAAGACGCGAGGACAACTTCCCTAC 3480







QY 1561 GCAAGGCGGCAACCCAGTGCACAGTCTGTGAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAGGCGGCAACCCAGTGCACAGTCTGTGAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCCTCAAGGATATTCAACAAGGGGTGAAGGATGCCAGAGGTACCCCATTTG 1680  
DB 1621 TGGCTCTCCTCAAGGATATTCAACAAGGGGTGAAGGATGCCAGAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGSCCTCGGTGCACATGCTTTATCATGTTTGTAGTCAGGTTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGSCCTCGGTGCACATGCTTTATCATGTTTGTAGTCAGGTTAAAA 1740  
QY 1741 AACGCTAGGCCCCCGAACCACAGGAGAGTGGTTTCTTTGAAACACGATAATACC 1800  
DB 1741 AACGCTAGGCCCCCGAACCACAGGAGAGTGGTTTCTTTGAAACACGATAATACC 1800  
QY 1801 ATGGCGCTATTAGCGGCTACTCCCAACAGACGGAGCGCTACTTGGCTGCATCACT 1860  
DB 1801 ATGGCGCTATTAGCGGCTACTCCCAACAGACGGAGCGCTACTTGGCTGCATCACT 1860  
QY 1861 AGCCTACAGCGCGGACAGAACAGGTCCAGGGGAGGTCCAAAGTGGTCTCCACCGCA 1920  
DB 1861 AGCCTACAGCGCGGACAGAACAGGTCCAGGGGAGGTCCAAAGTGGTCTCCACCGCA 1920  
QY 1921 ACACAAATCTTTCTGGGACCTCGGTCAATGGCGTGTGTGGACTGTCTATCATGTGCC 1980  
DB 1921 ACACAAATCTTTCTGGGACCTCGGTCAATGGCGTGTGTGGACTGTCTATCATGTGCC 1980  
QY 1981 GGCTCAAGACCTTTGGCGGCCAAAGGGCCCAATCAACCAATGTATGGAC 2040  
DB 1981 GGCTCAAGACCTTTGGCGGCCAAAGGGCCCAATCAACCAATGTATGGAC 2040  
QY 2041 CAGGACCTGTCGCTGCGACGCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2100  
DB 2041 CAGGACCTGTCGCTGCGACGCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2160  
DB 2101 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2160  
QY 2161 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2220  
DB 2161 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2220  
QY 2221 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2280  
DB 2221 GGCAGCTCGGACCTTACTTGGTCAAGGATGCCGATGCTTCCAGTGCACCTGC 2280  
QY 2281 ACCGAGGGGTTCGAGGGGTGGACTTTGTATCCCGTCCGAGTCTATGGAAACCACTATG 2340  
DB 2281 ACCGAGGGGTTCGAGGGGTGGACTTTGTATCCCGTCCGAGTCTATGGAAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCGCGCGTACCGAGACATTCAGGGT 2400  
DB 2341 CGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCGCGCGTACCGAGACATTCAGGGT 2400  
QY 2401 GCCCATCTACACGCCCTTACTGCTAGCGCAAGAGACATTAAGTGCAGGTCGGGTATGCA 2460  
DB 2401 GCCCATCTACACGCCCTTACTGCTAGCGCAAGAGACATTAAGTGCAGGTCGGGTATGCA 2460  
QY 2461 GCCCAAGGGTATTAAGGTGCTTGTCTTGAACCCCGTCCGCGCGCACCTTAGGTTTCGGG 2520  
DB 2461 GCCCAAGGGTATTAAGGTGCTTGTCTTGAACCCCGTCCGCGCGCACCTTAGGTTTCGGG 2520  
QY 2521 GCGTATATGCTTAAGGCACATGATATGACCCCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
DB 2521 GCGTATATGCTTAAGGCACATGATATGACCCCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
QY 2581 ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTCCGACGGTGTGTC 2640  
DB 2581 ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTCCGACGGTGTGTC 2640  
QY 2641 TCTGGGGGCGCCTATGACATCAATATGATGATGAGTGCCACTCACTGACTCGACCACT 2700

DB 2641 TCTGGGGGCGCCTATGACATCAATATGATGATGAGTGCCACTCACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCATTCGGCACAGTCTCTGGACCAACGAGAGACGGTGGAGCCGACTCTGTCGTG 2760  
DB 2701 ATCTGGGCATTCGGCACAGTCTCTGGACCAACGAGAGACGGTGGAGCCGACTCTGTCGTG 2760  
QY 2761 CTGCCACCGCTAGCCCTCGGGATTCGGTCACTCGGTGCACATCCAAACATCGAGAGGTG 2820  
DB 2761 CTGCCACCGCTAGCCCTCGGGATTCGGTCACTCGGTGCACATCCAAACATCGAGAGGTG 2820  
QY 2821 GCTCTGTCAGCACTGAGAAATCCCTTTATGGCAAGCAATCCCATCGACACATC 2880  
DB 2821 GCTCTGTCAGCACTGAGAAATCCCTTTATGGCAAGCAATCCCATCGACACATC 2880  
QY 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCCAAGAAAGAAATGTGATGAGCTCGCCGCG 2940  
DB 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCCAAGAAAGAAATGTGATGAGCTCGCCGCG 2940  
QY 2941 AAGCTGTCCGGCCTCGGACTCAATGCTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
DB 2941 AAGCTGTCCGGCCTCGGACTCAATGCTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
QY 3001 ATACCAACTAGCGAGAGCTCATTTGCTAGCAACGACGCTCTAAATGACGGGCTTTACC 3060  
DB 3001 ATACCAACTAGCGAGAGCTCATTTGCTAGCAACGACGCTCTAAATGACGGGCTTTACC 3060  
QY 3061 GGGGATTTGCACTCAGTGATCGACTGCAATACATGTGTCTACCCAGACAGTCACTCAGC 3120  
DB 3061 GGGGATTTGCACTCAGTGATCGACTGCAATACATGTGTCTACCCAGACAGTCACTCAGC 3120  
QY 3121 CTGGAACCGGACTTACCATTTGAGACGACGACGCTGCACAAGACGGGTGTACGCTCG 3180  
DB 3121 CTGGAACCGGACTTACCATTTGAGACGACGACGCTGCACAAGACGGGTGTACGCTCG 3180  
QY 3181 CAGCGGCGAGGAGGACTGCTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
DB 3181 CAGCGGCGAGGAGGACTGCTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATTTTCGATTTCTCGGTTCTGTGCGAGTGTATGACGGGGCTGT 3300  
DB 3241 GAACGGCCCTCGGGCATTTTCGATTTCTCGGTTCTGTGCGAGTGTATGACGGGGCTGT 3300  
QY 3301 GCTTGTGATCAGGCTACGCGCCGCGAGACCTCAGTTAGTGTGCGGCTTACCTAAACACA 3360  
DB 3301 GCTTGTGATCAGGCTACGCGCCGCGAGACCTCAGTTAGTGTGCGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCGGCTGCGGAGGACCTCTGGAGTCTGGAGAGGCTCTTTACAGGCTC 3420  
DB 3361 CCAGGTTGCGGCTGCGGAGGACCTCTGGAGTCTGGAGAGGCTCTTTACAGGCTC 3420  
QY 3421 ACCCATATAGACCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTAC 3480  
DB 3421 ACCCATATAGACCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTAC 3480  
QY 3481 CTGGTATGATACAGGCTACGCTGTGGGAGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
DB 3481 CTGGTATGATACAGGCTACGCTGTGGGAGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
QY 3541 CAAATCTGGAAGTCTCATAGGCTAAAGCTTACGCTGCAAGGCGCAAGCCCTGCTG 3600  
DB 3541 CAAATCTGGAAGTCTCATAGGCTAAAGCTTACGCTGCAAGGCGCAAGCCCTGCTG 3600  
QY 3601 TATAGCTGGGAGCCGTTTCAAAACAGAGTTTACTTACCACACACCCCAATACCAATACATC 3660  
DB 3601 TATAGCTGGGAGCCGTTTCAAAACAGAGTTTACTTACCACACACCCCAATACCAATACATC 3660  
QY 3661 ATGGCATGATGCTCGCTCAGCTTGGAGGTGCTCAGAGACCTTGGGTGCTGTAGCGGA 3720  
DB 3661 ATGGCATGATGCTCGCTCAGCTTGGAGGTGCTCAGAGACCTTGGGTGCTGTAGCGGA 3720  
QY 3721 GTCCCTAGACCTCTGCGCCGCTTATGCTCTGACAAAGGAGGCTGTGCTATTTGGGCGAG 3780



Db 3721 GTCCTAGCAGCTCGGCCGCTATTGCTGACAA CAGGACAGCGTGGTCAATCTGCGGAGG 3780  
QY 3781 ATCATCTTTGTCGGAAAGCGGCCATCATTTCCGGA CAGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTTGTCGGAAAGCGGCCATCATTTCCGGA CAGGAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTATCGA AAGGAAATGCAAGTCCGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTATCGA AAGGAAATGCAAGTCCGCC 3900  
QY 3901 GAAACAATTTCAACAGAGGCAATTCGGGTGCTGCAAA CAGCCACCAAGCAGCGAGGCT 3960  
Db 3901 GAACAAATTTCAACAGAGGCAATTCGGGTGCTGCAAA CAGCCACCAAGCAGCGAGGCT 3960  
QY 3961 GTGTCTCCGCTGGTGAATCCAAAGTGGCGGAACCTCGA AAGCCTTCTGGCGGAAGCATATG 4020  
Db 3961 GTGTCTCCGCTGGTGAATCCAAAGTGGCGGAACCTCGA AAGCCTTCTGGCGGAAGCATATG 4020  
QY 4021 TGGAAATTTTATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTTTATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGCAATTCAGAGCTCTATCACCA GCGCGGTCAACCAACCAAT 4140  
Db 4081 GCGATAGCATCACTGATGCAATTCAGAGCTCTATCACCA GCGCGGTCAACCAACCAAT 4140  
QY 4141 ACCCTCTCTGTTAAACATCTGGGGGATGGGTGGCGGCCA ACTTGTCTCTCCGAGCGT 4200  
Db 4141 ACCCTCTCTGTTAAACATCTGGGGGATGGGTGGCGGCCA ACTTGTCTCTCCGAGCGT 4200  
QY 4201 GCTTCTGCTTTTAAACATCTGGGGGATGGGTGGCGGCCA ACTTGTCTCTCCGAGCGT 4260  
Db 4201 GCTTCTGCTTTTAAACATCTGGGGGATGGGTGGCGGCCA ACTTGTCTCTCCGAGCGT 4260  
QY 4261 AAGGTGCTTTGATATTTTGGCAGGTATGGAGCAGGGGTGG CAGCGCGTCTGTTGGCC 4320  
Db 4261 AAGGTGCTTTGATATTTTGGCAGGTATGGAGCAGGGGTGG CAGCGCGTCTGTTGGCC 4320  
QY 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCGACGAGGACTGG TTAACCTACTCTCCCTGCT 4380  
Db 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCGACGAGGACTGG TTAACCTACTCTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCGTCGGGTCTGTGCGC AGCGATCTGCGTCCGAC 4440  
Db 4381 ATCTCTCTCCCTGGCGCCCTAGTCGTCGGGTCTGTGCGC AGCGATCTGCGTCCGAC 4440  
QY 4441 GTGGGCCACAGGGAGGGGGCTGTGCAGTGAATGAACCGG CTGATAGCGTTCGCTCCGCG 4500  
Db 4441 GTGGGCCACAGGGAGGGGGCTGTGCAGTGAATGAACCGG CTGATAGCGTTCGCTCCGCG 4500  
QY 4501 GGTAAACCAAGTCTCCCGCAGCACTATGTGCTGAGAGGAGCG CTGAGCAAGTGTCACT 4560  
Db 4501 GGTAAACCAAGTCTCCCGCAGCACTATGTGCTGAGAGGAGCG CTGAGCAAGTGTCACT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGTCTCTGAAGG GTTCCACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGTCTCTGAAGG GTTCCACAGTGGATCAAC 4620  
QY 4621 GAGGACTGTCTCCAGCCATCTCCGCTCGTGGCTAAGAGA TGTTTGGGATGGATATGC 4680  
Db 4621 GAGGACTGTCTCCAGCCATCTCCGCTCGTGGCTAAGAGA TGTTTGGGATGGATATGC 4680  
QY 4681 ACGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAGTCCAG TCCCTGCGGAGTTCGCGGA 4740  
Db 4681 ACGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAGTCCAG TCCCTGCGGAGTTCGCGGA 4740  
QY 4741 GTCCCTCTCTCTCATGTCAGTGAAGTGGGTCAAGGGAGTCT GCGGGCGGACGCAATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCAGTGAAGTGGGTCAAGGGAGTCT GCGGGCGGACGCAATCATG 4800  
QY 4801 CAAACCACTCTGCCATGTGGAGCA CAGATCACCGGACATGTGAAAAACGTTTCCATGAGG 4860  
Db 4801 CAAACCACTCTGCCATGTGGAGCA CAGATCACCGGACATGTGAAAAACGTTTCCATGAGG 4860

QY 4861 ATCTGGGGCCCTAGGACCTGTAGTAACACGTGGCATGGAACAT TCCCATTAAACCGGTAC 4920  
Db 4861 ATCTGGGGCCCTAGGACCTGTAGTAACACGTGGCATGGAACAT TCCCATTAAACCGGTAC 4920  
QY 4921 ACCAGGGCCCTGACGCGCTCCCGCGCCAAATTTATTTAG GGGCGTGTGGGGGTG 4980  
Db 4921 ACCAGGGCCCTGACGCGCTCCCGCGCCAAATTTATTTAG GGGCGTGTGGGGGTG 4980  
QY 4981 GCTGTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCAT CAGTGTGACGGGCATG 5040  
Db 4981 GCTGTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCAT CAGTGTGACGGGCATG 5040  
QY 5041 ACCACTGACAAACGTAAAGTGGCCGCTCAGGTTCCGGGCC CCGAAATTTCTTCAAGAAATG 5100  
Db 5041 ACCACTGACAAACGTAAAGTGGCCGCTCAGGTTCCGGGCC CCGAAATTTCTTCAAGAAATG 5100  
QY 5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCC CTCTACGGGAGAGTTC 5160  
Db 5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCC CTCTACGGGAGAGTTC 5160  
QY 5161 ACATTTCTTGGTGGGCTCAATCAATACCTGTTGGTCAAGTCC CAGTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTTCTTGGTGGGCTCAATCAATACCTGTTGGTCAAGTCC CAGTCCCATGCGAGCCGAA 5220  
QY 5221 CCGGACGTAGCAGTCTCACTTCCATGCTCACCGACCCCTCC CAATTAACGGCGGAGAG 5280  
Db 5221 CCGGACGTAGCAGTCTCACTTCCATGCTCACCGACCCCTCC CAATTAACGGCGGAGAG 5280  
QY 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGG CCAGGTCAATCAGTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTTGG CCAGGTCAATCAGTAGCCAG 5340  
QY 5341 CTGTCTGGCCTTCTTGAAGCAACATGCACTACCCGTCAATG CACTCCCGGAGCGTAC 5400  
Db 5341 CTGTCTGGCCTTCTTGAAGCAACATGCACTACCCGTCAATG CACTCCCGGAGCGTAC 5400  
QY 5401 CTCATCGAGGCCAACTCTGTGGGGCAGGAGATGGCGGGAAC ATCAACCCGCTGGAG 5460  
Db 5401 CTCATCGAGGCCAACTCTGTGGGGCAGGAGATGGCGGGAAC ATCAACCCGCTGGAG 5460  
QY 5461 TCAGAAAATAAGGTAGTAATTTTGAATCTTTTGGAGTCTTT TCGAGCCGCTCCAGCGAGGAGTATG 5520  
Db 5461 TCAGAAAATAAGGTAGTAATTTTGAATCTTTTGGAGTCTTT TCGAGCCGCTCCAGCGAGGAGTATG 5520  
QY 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTTTCGAGAGGTCCAG GAAATTTCCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTTTCGAGAGGTCCAG GAAATTTCCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGCACGCGCGGATTAACACCTCTTACAGTCTTGA GAGTCTTGAAGGACCCGGAC 5640  
Db 5581 CCCATATGGGCACGCGCGGATTAACACCTCTTACAGTCTTGA GAGTCTTGAAGGACCCGGAC 5640  
QY 5641 TAGTCCCTCCAGTGGTACACGGGTTCATTGCGCCCTGCGAAG CCGCTCCGATACCA 5700  
Db 5641 TAGTCCCTCCAGTGGTACACGGGTTCATTGCGCCCTGCGAAG CCGCTCCGATACCA 5700  
QY 5701 CCTTCCAGGAGNAGGAGCGGTTGCTGTGAGATCTACCGTGTCT CCGTCTTGGCG 5760  
Db 5701 CCTTCCAGGAGNAGGAGCGGTTGCTGTGAGATCTACCGTGTCT CCGTCTTGGCG 5760  
QY 5761 GAGCTGCCACAAAGACCTTTCGCGAGTCCGAAATCTGCGCCGT TCGACAGCGGACCGCA 5820  
Db 5761 GAGCTGCCACAAAGACCTTTCGCGAGTCCGAAATCTGCGCCGT TCGACAGCGGACCGCA 5820  
QY 5821 ACGCCCTCTCTTACAGAGCCCTCCGACAGCGGAGTCCGAGAT TCGAGCGTGTAGTCTG 5880  
Db 5821 ACGCCCTCTCTTACAGAGCCCTCCGACAGCGGAGTCCGAGAT TCGAGCGTGTAGTCTG 5880  
QY 5881 TCTTCCATGCCCCCTTTCAGGGGAGCGGGGATCCCGATCTTCA GCGAGCGGTCTTGG 5940  
Db 5881 TCTTCCATGCCCCCTTTCAGGGGAGCGGGGATCCCGATCTTCA GCGAGCGGTCTTGG 5940





; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION  
 ; FILE REFERENCE: 0342/1H395US3  
 ; CURRENT APPLICATION NUMBER: US 10/434,842  
 ; CURRENT FILING DATE: 2003-05-09  
 ; PRIOR APPLICATION NUMBER: US 10/233,307  
 ; PRIOR FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: US 10/005,469  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV9 subgenomic HCV replicon  
 US-10-434-842-4

Query Match 99.9%; Score 7981; DB 16; Length 7992;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7984; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	GCACGCCCGATTGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAACACTG	60
DB	1	GCACGCCCGATTGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAACACTG	60
QY	61	TCTTACGAGAGAGCGTCTAGCCATGGCGTTAGTATGATGTCGTCGAGCCTCCAGGAC	120
DB	61	TCTTACGAGAGAGCGTCTAGCCATGGCGTTAGTATGATGTCGTCGAGCCTCCAGGAC	120
QY	121	CCCCCTCCCGGAGAGCGATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCGAG	180
DB	121	CCCCCTCCCGGAGAGCGATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCGAG	180
QY	181	GAGACCGGGTCTTTCTGGATCAACCGCTCAATGCCCTGGAGATTGGGGCTGCCCC	240
DB	181	GAGACCGGGTCTTTCTGGATCAACCGCTCAATGCCCTGGAGATTGGGGCTGCCCC	240
QY	241	GCAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCGCTTGTGCTACTGCTGATAGG	300
DB	241	GCAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCGCTTGTGCTACTGCTGATAGG	300
QY	301	GTGCTTCGAGTGCCTCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAC	360
DB	301	GTGCTTCGAGTGCCTCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAC	360
QY	361	CTCAAGAAACCAAGGGCGCCCATGATTGAACAAGATGATTGACGCGAGTTCTC	420
DB	361	CTCAAGAAACCAAGGGCGCCCATGATTGAACAAGATGATTGACGCGAGTTCTC	420
QY	421	CGCCCGCTTGGTGGAGAGCTATTGCGCTATGATGGGCACACAGCAATCGGTGCT	480
DB	421	CGCCCGCTTGGTGGAGAGCTATTGCGCTATGATGGGCACACAGCAATCGGTGCT	480
QY	481	CTGATGCGCGGTGTTCCGCTGTACGCGAGGGGCGCGGTCTTTTGTCAAGACCG	540
DB	481	CTGATGCGCGGTGTTCCGCTGTACGCGAGGGGCGCGGTCTTTTGTCAAGACCG	540
QY	541	ACCTGTCGCGTGCCTCAATGAATCTGAGGACGAGGCGCGCTATCGTGGCTGGCCA	600
DB	541	ACCTGTCGCGTGCCTCAATGAATCTGAGGACGAGGCGCGCTATCGTGGCTGGCCA	600
QY	601	CGACGGCGTTCCTTGGCGAGCTGTGCTGAGCGTTGTCACTGAAGCGGGAAGGACTGGC	660
DB	601	CGACGGCGTTCCTTGGCGAGCTGTGCTGAGCGTTGTCACTGAAGCGGGAAGGACTGGC	660
QY	661	TGCTATTGGGGAAGTCCGGGGCAGGATCTCCTGTCACTCACTTGTCTCTGCCGAGA	720
DB	661	TGCTATTGGGGAAGTCCGGGGCAGGATCTCCTGTCACTCACTTGTCTCTGCCGAGA	720

QY	721	AAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATTCGGCTACTGCTGC	780
DB	721	AAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATTCGGCTACTGCTGC	780
QY	781	CATTGACACCAAGGAAACATCGCATCGAGGACGACTACTCGGATGGAAGCGGTC	840
DB	781	CATTGACACCAAGGAAACATCGCATCGAGGACGACTACTCGGATGGAAGCGGTC	840
QY	841	TTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGAATCTGTCG	900
DB	841	TTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGAATCTGTCG	900
QY	901	CCAGGCTCAAGCGCGCATGCGCGAGAGATCTCGTGTGACCCATGCGATGCGCT	960
DB	901	CCAGGCTCAAGCGCGCATGCGCGAGAGATCTCGTGTGACCCATGCGATGCGCT	960
QY	961	GCTTGGCGAATATCATGTGGAAATGGCGCTTTTCTGGATTCATCGATCTGGCGGC	1020
DB	961	GCTTGGCGAATATCATGTGGAAATGGCGCTTTTCTGGATTCATCGATCTGGCGGC	1020
QY	1021	TGGGTGTGGCGACGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGC	1080
DB	1021	TGGGTGTGGCGACGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGC	1080
QY	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTTCACCGTATCGCGCTCCCGATTGCG	1140
DB	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTTCACCGTATCGCGCTCCCGATTGCG	1140
QY	1141	AGGGATCGCTTCTATCGCTTCTTGAGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
DB	1141	AGGGATCGCTTCTATCGCTTCTTGAGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
QY	1201	GTTCCTCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAACGTTACTGCG	1260
DB	1201	GTTCCTCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAACGTTACTGCG	1260
QY	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCTTGTCTATATGTTATTTTCCACATATTG	1320
DB	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCTTGTCTATATGTTATTTTCCACATATTG	1320
QY	1321	CGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTCTCTTCTTGAGGACATTCCT	1380
DB	1321	CGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTCTCTTCTTGAGGACATTCCT	1380
QY	1381	AGGGGTCTTTTCCCTCTCGCAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAGCA	1440
DB	1381	AGGGGTCTTTTCCCTCTCGCAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAGCA	1440
QY	1441	GTTCCTCTGGAAGCTTTTGAAGACAAACAAACGCTGTAGCGACCTTTTGCAGGCGCG	1500
DB	1441	GTTCCTCTGGAAGCTTTTGAAGACAAACAAACGCTGTAGCGACCTTTTGCAGGCGCG	1500
QY	1501	AACCCCGACCTGGCGACAGGTGCTTCTGGCGCAAGGCGGCTGTATGAAGATACACCT	1560
DB	1501	AACCCCGACCTGGCGACAGGTGCTTCTGGCGCAAGGCGGCTGTATGAAGATACACCT	1560
QY	1561	GCAAGCGCGCAACACCCAGTGCACGTTGTGAGTTGATGTTGTGAAGAGAGTCAAA	1620
DB	1561	GCAAGCGCGCAACACCCAGTGCACGTTGTGAGTTGATGTTGTGAAGAGAGTCAAA	1620
QY	1621	TGCTCTCTCTCAAGCGTATTCACCAAGGGGCTGAAGGATGCGGAGGCTTAAATTT	1680
DB	1621	TGCTCTCTCTCAAGCGTATTCACCAAGGGGCTGAAGGATGCGGAGGCTTAAATTT	1680
QY	1681	ATGGGATCTGATCTGGGGCTCGGTGCATGCTTTACATGCTTTAGTGTAGGTTAAAA	1740
DB	1681	ATGGGATCTGATCTGGGGCTCGGTGCATGCTTTACATGCTTTAGTGTAGGTTAAAA	1740
QY	1741	AACGCTCTAGGCGCCCGGAAACACGCGGACGTTGTTTCTTTTGAAGAAACAGTAAATCC	1800
DB	1741	AACGCTCTAGGCGCCCGGAAACACGCGGACGTTGTTTCTTTTGAAGAAACAGTAAATCC	1800
QY	1801	ATGGCGCTATTACGGCTTACTCCCAACAGACGGGCGCTACTTGGCTGCATCACT	1860

[illegible]

Db	2881	AAGGGGGGAGGACCTCAITTTCTGCCAATTC	CAAGAAAGAAATGTGATGAGCTGCCGCG	2941
Qy	2941	AAGCTGTCCGCTCGGACTCAATGCTGTAGCA	TATTAACGGGGCCCTTGATGATCCGCT	3001
Db	2941			3001
Db	2941	AAGCTGTCCGCTCGGACTCAATGCTGTAGCA	TATTAACGGGGCCCTTGATGATCCGCT	3000
Qy	3001	ATACCAACTAGCGGAGAGCTCAITTTGCTG	TAGCAACGGACGCTCTAAATGACGGGCTTTACC	3060
Db	3001	ATACCAACTAGCGGAGAGCTCAITTTGCTG	TAGCAACGGACGCTCTAATGACGGGCTTTACC	3060
Qy	3061	GGCGATTTGCACTCAGTGATCGACTGCAAT	ACATATGTTGTCAACCGACAGTCGACTTCAGC	3120
Db	3061	GGCGATTTGCACTCAGTGATCGACTGCAAT	ACATATGTTGTCAACCGACAGTCGACTTCAGC	3120
Qy	3121	CTGACCCGACCTTCACTTTGCTGAGACGAC	CGCGTGCCACAAGACGGGTGTCAAGCTCG	3180
Db	3121	CTGACCCGACCTTCACTTTGCTGAGACGAC	CGCGTGCCACAAGACGGGTGTCAAGCTCG	3180
Qy	3181	CAGCGGAGGACGAGCACTGGTAGGGGAGG	ATGGGCATTTTACAGGTTTGTGATCTCCAGGA	3240
Db	3181	CAGCGGAGGACGAGCACTGGTAGGGGAGG	ATGGGCATTTTACAGGTTTGTGATCTCCAGGA	3240
Qy	3241	GAACGGCCCTCGGCGCATGTTCTCGGTTCT	GTGCGAGTGCTATGACGGGGCTGT	3300
Db	3241	GAACGGCCCTCGGCGCATGTTCTCGGTTCT	GTGCGAGTGCTATGACGGGGCTGT	3300
Qy	3301	GCTTTGTACGAGCTCACGCCGCGGACACT	CTAGTTAGGTTGCGGCTTACCTAAACACA	3360
Db	3301	GCTTTGTACGAGCTCACGCCGCGGACACT	CTAGTTAGGTTGCGGCTTACCTAAACACA	3360
Qy	3361	CCAGGTTGCCCGTCTGCCAGGACCATCTG	GAGTTCTGGGAGAGCGTCTTACAGGCTTC	3420
Db	3361	CCAGGTTGCCCGTCTGCCAGGACCATCTG	GAGTTCTGGGAGAGCGTCTTACAGGCTTC	3420
Qy	3421	ACCACATAGACGCCCATTTCTTGTCAGACT	TAAGCAGCAGGAGACAATTCCTCCCTAC	3480
Db	3421	ACCACATAGACGCCCATTTCTTGTCAGACT	TAAGCAGCAGGAGACAATTCCTCCCTAC	3480
Qy	3481	CTGCTAGCATACAGGCTACGGTGTGGCC	CAGGGCTCAGGCTCCACTCCATCGTGGAC	3540
Db	3481	CTGCTAGCATACAGGCTACGGTGTGGCC	CAGGGCTCAGGCTCCACTCCATCGTGGAC	3540
Qy	3541	CAAAATGTGAAGTGTCTCATACGGCTTAA	AGCCTACGCTGACCGGGCCACGCCCTGCTG	3600
Db	3541	CAAAATGTGAAGTGTCTCATACGGCTTAA	AGCCTACGCTGACCGGGCCACGCCCTGCTG	3600
Qy	3601	TATAGGCTGGAGCCGTTCAAAACGAGGTT	ATTACACACACCCCAATACATC	3660
Db	3601	TATAGGCTGGAGCCGTTCAAAACGAGGTT	ATTACACACACCCCAATACATC	3660
Qy	3661	ATGGCATGCAATGCTGGCTGACCTGAGG	TGCTTACGAGCACCTGGGTGCTGTAGCGGA	3720
Db	3661	ATGGCATGCAATGCTGGCTGACCTGAGG	TGCTTACGAGCACCTGGGTGCTGTAGCGGA	3720
Qy	3721	GTCCTAGCAGCTGTGGCCGGTATTGCTT	GACAAACAGGAGCGTGTCTATTGTGGCAGG	3780
Db	3721	GTCCTAGCAGCTGTGGCCGGTATTGCTT	GACAAACAGGAGCGTGTCTATTGTGGCAGG	3780
Qy	3781	ATCATTTTGTCCGGAAGCGGCGCATCAT	TCCCGACAGGGAAGTCTTTACCGGGAGTTC	3840
Db	3781	ATCATTTTGTCCGGAAGCGGCGCATCAT	TCCCGACAGGGAAGTCTTTACCGGGAGTTC	3840
Qy	3841	GATGAGATGAAGAGTGGCCCTCACACCT	CTCCCTTACATCGAACAGGGAATGAGCTCGCC	3900
Db	3841	GATGAGATGAAGAGTGGCCCTCACACCT	CTCCCTTACATCGAACAGGGAATGAGCTCGCC	3900
Qy	3901	GAACAAATTCAAACAGAAAGCAATCGG	GTGCTGCAAAACAGCCAAAGCAGCGAGGCT	3960
Db	3901	GAACAAATTCAAACAGAAAGCAATCGG	GTGCTGCAAAACAGCCAAAGCAGCGAGGCT	3960
Qy	3961	GCTGCTCCCGTGGTGAATTCCAAGTGG	CGGACCTCTCGAAGCCCTTTCTGGCGAAGCATATG	4020
Db	3961	GCTGCTCCCGTGGTGAATTCCAAGTGG	CGGACCTCTCGAAGCCCTTTCTGGCGAAGCATATG	4020

QY	4021	TGGAATTTTCATCAGCGGGATACAATAATTTAGCAGGCTTGTCCA	CTGTGCTGGCAACCCC	4080
Db	4021	TGGAATTTTCATCAGCGGGATACAATAATTTAGCAGGCTTGTCCA	CTGTGCTGGCAACCCC	4080
QY	4081	GGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGACCGCT	CACCAACCCACAT	4140
Db	4081	GGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGACCGCT	CACCAACCCACAT	4140
QY	4141	ACCCTCGTGTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTG	CTCTCTCCAGCGCT	4200
Db	4141	ACCCTCGTGTAAACATCTCTGGGGGATGGTGGCGGCCCAACTTG	CTCTCTCCAGCGCT	4200
QY	4201	GCTTCTGCTTTCGTAGCGCGCGGCATCTGTGAGACGGCTGT	TGGCAGCATPAGGCTTGGG	4260
Db	4201	GCTTCTGCTTTCGTAGCGCGCGGCATCTGTGAGACGGCTGT	TGGCAGCATPAGGCTTGGG	4260
QY	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGCA	GCAGCGCGCTCTGTGGCC	4320
Db	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGCA	GCAGCGCGCTCTGTGGCC	4320
QY	4321	TTTAAAGTCAATGAGCGCGAGATGCGCCTCCACCGAGACCTGTG	TAACTACTCCCTGCT	4380
Db	4321	TTTAAAGTCAATGAGCGCGAGATGCGCCTCCACCGAGACCTGTG	TAACTACTCCCTGCT	4380
QY	4381	ATCCTCTCCCTCGCGCCCTAGTCTGTGCGGGTCTGTGCGCAGCG	ACTACTGCTCGGCAC	4440
Db	4381	ATCCTCTCCCTCGCGCCCTAGTCTGTGCGGGTCTGTGCGCAGCG	ACTACTGCTCGGCAC	4440
QY	4441	GTGGGCCAAGGAGGGGGCTGTGCATGTGATAAACCGCTGTATAG	CGTTTCGGTTCGCGG	4500
Db	4441	GTGGGCCAAGGAGGGGGCTGTGCATGTGATAAACCGCTGTATAG	CGTTTCGGTTCGCGG	4500
QY	4501	GGTAACCAAGTCTCCCCCA	CGCATATGTGCTGTGAGACGACGCTGCAGCAGTGTCACT	4560
Db	4501	GGTAACCAAGTCTCCCCCA	CGCATATGTGCTGTGAGACGACGCTGCAGCAGTGTCACT	4560
QY	4561	CAGATCCTCTAGTCTTACCATCACTCAGTCTGTGAAGAGGCTT	CACCAAGTGGATCAAC	4620
Db	4561	CAGATCCTCTAGTCTTACCATCACTCAGTCTGTGAAGAGGCTT	CACCAAGTGGATCAAC	4620
QY	4621	GAGGACTGTCTCCAGCCCATCTCCGGCTCGTGGCTAAGAGATG	TTTGGGATTCGATATGC	4680
Db	4621	GAGGACTGTCTCCAGCCCATCTCCGGCTCGTGGCTAAGAGATG	TTTGGGATTCGATATGC	4680
QY	4681	ACGGTGTAGTCAATTCACAGACCTGCTCCAGTCCAGCTCTCTG	CCGCGGATTCGCGGGA	4740
Db	4681	ACGGTGTAGTCAATTCACAGACCTGCTCCAGTCCAGCTCTCTG	CCGCGGATTCGCGGGA	4740
QY	4741	GTCCCTCTCTCTCATGTCAACGTGGGTACAAAGGAGTCTGGG	GGGCGCAGCGCATCATG	4800
Db	4741	GTCCCTCTCTCTCATGTCAACGTGGGTACAAAGGAGTCTGGG	GGGCGCAGCGCATCATG	4800
QY	4801	CAAAACACTTGCCTATGTGGAGACAGATCACCGGACATGTG	TAATAAACGGTTCATGAGG	4860
Db	4801	CAAAACACTTGCCTATGTGGAGACAGATCACCGGACATGTG	TAATAAACGGTTCATGAGG	4860
QY	4861	ATCTGGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACAT	TTCGCCATTAACCGGTAC	4920
Db	4861	ATCTGGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACAT	TTCGCCATTAACCGGTAC	4920
QY	4921	ACCAACGGGCCCCGTGACAGCCCTCCCGCGCCCAAAATTTAT	TCTAGGGCGCTGTGCGGGTG	4980
Db	4921	ACCAACGGGCCCCGTGACAGCCCTCCCGCGCCCAAAATTTAT	TCTAGGGCGCTGTGCGGGTG	4980
QY	4981	GCTGCTCAGGAGTACGTGGAGTTTACCGGGTGGGGATTTT	CCACTACTGTCAGCGGCATG	5040
Db	4981	GCTGCTCAGGAGTACGTGGAGTTTACCGGGTGGGGATTTT	CCACTACTGTCAGCGGCATG	5040
QY	5041	ACCACTGACAACTGTAAGTGCCTGTTCAGGTTCCGGCCCCCG	GAATTCCTTACAGAAAGTG	5100
Db	5041	ACCACTGACAACTGTAAGTGCCTGTTCAGGTTCCGGCCCCCG	GAATTCCTTACAGAAAGTG	5100

Qy	5101	GATGGGTTGGGTTGACACAGGTACGCTCCAGCGTCAAAACCCCTCCTTACGGGAGGAGTC	5161
Db	5101	GATGGGTTGGGTTGACACAGGTACGCTCCAGCGTCAAAACCCCTCCTTACGGGAGGAGTC	5160
Qy	5161	ACATTCTCGTGGGCTCAATCAATACCTGTGTTGGTTCACAGCTCCCATGCGAGCCCGAA	5220
Db	5161	ACATTCTCGTGGGCTCAATCAATACCTGTGTTGGTTCACAGCTCCCATGCGAGCCCGAA	5220
Qy	5221	CCGGAGTACAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTAACGGCGGAGACG	5280
Db	5221	CCGGAGTACAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTAACGGCGGAGACG	5280
Qy	5281	GCTAAGGTAGGTGCGCAGGGGATCTCCCCCTCTTGGCCAGCTCATCAGTAGTACCCAG	5340
Db	5281	GCTAAGGTAGGTGCGCAGGGGATCTCCCCCTCTTGGCCAGCTCATCAGTAGTACCCAG	5340
Qy	5341	CTGCTCGGCCCTTCTTTGAAGGCACATGACACTACCCGTTCATCACTCCCCCGGACGCTGAC	5400
Db	5341	CTGCTCGGCCCTTCTTTGAAGGCACATGACACTACCCGTTCATCACTCCCCCGGACGCTGAC	5400
Qy	5401	CTCATCGAGCCCAACCTCTGTGGCGCAGGAGATGGGCGGGAACATCAACCCGCTGGAG	5460
Db	5401	CTCATCGAGCCCAACCTCTGTGGCGCAGGAGATGGGCGGGAACATCAACCCGCTGGAG	5460
Qy	5461	TCAGAAAATAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAAGCGGAGGAGATGAG	5520
Db	5461	TCAGAAAATAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAAGCGGAGGAGATGAG	5520
Qy	5521	AGGAGTATCCCTTCGGCGGAGATCCTCGCGAGGTCAGAGGAAATTCCTCTCGAGCCATG	5580
Db	5521	AGGAGTATCCCTTCGGCGGAGATCCTCGCGAGGTCAGAGGAAATTCCTCTCGAGCCATG	5580
Qy	5581	CCCATATGGGCACGCGCCGGAATTACACCCCTCCACTGTTAGAGTCTTGGAAGGACCCGGAC	5640
Db	5581	CCCATATGGGCACGCGCCGGAATTACACCCCTCCACTGTTAGAGTCTTGGAAGGACCCGGAC	5640
Qy	5641	TACGTCCCTCCAGTGGTACACGSGTGTCAATGCGCGCTCGCAAGGCCCTCCGATACCA	5700
Db	5641	TACGTCCCTCCAGTGGTACACGSGTGTCAATGCGCGCTCGCAAGGCCCTCCGATACCA	5700
Qy	5701	CCTCCAGGAGGAGGAGCGTTGCTCTGTACAGATCTACCGTGCTTCGCTCTGGCG	5760
Db	5701	CCTCCAGGAGGAGGAGCGTTGCTCTGTACAGATCTACCGTGCTTCGCTCTGGCG	5760
Qy	5761	GAGCTCCGCAAAAGACCTTCGGCAGCTCCGGAATCGTCGCGCGCTCGACAGCGCACGGCA	5820
Db	5761	GAGCTCCGCAAAAGACCTTCGGCAGCTCCGGAATCGTCGCGCGCTCGACAGCGCACGGCA	5820
Qy	5821	ACGCGCTCTCTGACAGCCCTCCGACGACGGCGACGCGGATCCGAGCTTGAGTCTGATC	5880
Db	5821	ACGCGCTCTCTGACAGCCCTCCGACGACGGCGACGCGGATCCGAGCTTGAGTCTGATC	5880
Qy	5881	TCCTCATGCCCCCTTGAGGGGAGCCCGGGATCCCGATCTCAGCGAGCGGTCCTGG	5940
Db	5881	TCCTCATGCCCCCTTGAGGGGAGCCCGGGATCCCGATCTCAGCGAGCGGTCCTGG	5940
Qy	5941	TCTACCGTAAGCAGGAGGCTAGTGAGGACGTGCTCTGCTCGATGTCTACACATGG	6000
Db	5941	TCTACCGTAAGCAGGAGGCTAGTGAGGACGTGCTCTGCTCGATGTCTACACATGG	6000
Qy	6001	ACAGGCGCCTGATCACGCCATCGCTGCGGAGGAAACCAAGTGTGCCATCAATGCACTG	6060
Db	6001	ACAGGCGCCTGATCACGCCATCGCTGCGGAGGAAACCAAGTGTGCCATCAATGCACTG	6060
Qy	6061	AGCAACTCTTGTCTCGTCAACCAACTTGGTCTATGCTACACATCTCCGAGCGCAAGC	6120
Db	6061	AGCAACTCTTGTCTCGTCAACCAACTTGGTCTATGCTACACATCTCCGAGCGCAAGC	6120
Qy	6121	CTCGGCGAGGAAGGTCACTTTGACAGACTGCAAGTCTCTGGACGACCACTACCGGAC	6180
Db	6121	CTCGGCGAGGAAGGTCACTTTGACAGACTGCAAGTCTCTGGACGACCACTACCGGAC	6180
Qy	6181	GTGCTCAAGAGATGAAGGCGAAGGCGCTCCACAGTTAAGGCTAAACTTCTATCCGTGGAG	6240



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:11:15 ; Search time 21853 Seconds

(without alignments)  
13321.604 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccgatggggc.....ggcctctcagatcaagt 7989

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_est8.\*

9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795.4	10.0	935	7	CK284786 EST747508
2	794	9.9	804	7	CK291519 EST754233
3	794	9.9	856	7	CK287297 EST750019
4	794	9.9	910	7	CK287930 EST750652
5	794	9.9	933	7	CK291799 EST754513
6	794	9.9	936	7	CK256977 EST740614
7	794	9.9	947	7	CK298208 EST760922
8	794	9.9	954	7	CK283361 EST746083
9	778.8	9.7	811	7	CK288185 EST750907
10	587	7.3	878	7	CK288711 EST751433
11	577.4	7.2	789	8	AQ361914 mgxb0005K
12	559.8	7.0	620	8	AQ398387 mgxb0009J
13	542.4	6.8	549	7	CK801630 NF32c08f4
14	514.8	6.4	561	8	AQ447874 mgxb0012I
15	508	6.4	509	8	AQ447775 mgxb0011E
16	505.4	6.3	637	8	AQ47230 mgxb0006O
17	503.8	6.3	592	8	AQ49162 mgxb0023E
18	499	6.2	499	8	AQ398880 mgxb0006P
19	496.4	6.2	498	8	AQ398730 mgxb0009G
20	477.4	6.0	586	7	CK801096 NF15h06f4
21	474.8	6.0	857	8	AQ875013 V120B mt
22	474.8	5.9	710	8	BH972732 odj46t10
23	469.8	5.9	473	8	AQ397768 mgxb0001J
24	468.8	5.9	594	8	AQ397253 mgxb0003C

C 25	465.2	5.8	850	8	AQ875052	AQ875052 V120F5 mt
C 26	463.4	5.8	798	8	AQ876139	AQ876139 V133G10 m
C 27	458.8	5.7	482	1	AL449872	AL449872 AL449872
C 28	456.8	5.7	861	8	AQ875040	AQ875040 V120E4 mt
C 29	455.6	5.7	815	8	AQ876134	AQ876134 V133F5 mt
C 30	451.2	5.6	799	8	AQ876220	AQ876220 V152G7 mt
C 31	450	5.6	784	8	AQ876152	AQ876152 V133H12 m
C 32	444.6	5.6	790	8	AQ875904	AQ875904 V130H11 m
C 33	437.2	5.5	797	8	AQ876297	AQ876297 V153G3 mt
C 34	434.6	5.4	765	7	CK283362	CK283362 EST746084
C 35	426.8	5.3	804	8	AQ876080	AQ876080 V133A4 mt
C 36	422.8	5.3	759	7	CK298209	CK298209 EST760923
C 37	414.6	5.2	791	8	AQ876019	AQ876019 V132C3 mt
C 38	409.4	5.1	748	8	AQ501530	AQ501530 V20F6 mtH
C 39	408	5.1	635	8	AQ447643	AQ447643 mgxb0008M
C 40	408	5.1	656	8	AQ447140	AQ447140 mgxb0005I
C 41	404	5.1	790	8	AQ875912	AQ875912 V130H8 mt
C 42	404	5.1	791	8	AQ876201	AQ876201 V152E2 mt
C 43	402.6	5.0	856	8	AQ875050	AQ875050 V120F2 mt
C 44	401.8	5.0	760	7	CK287931	CK287931 EST750653
C 45	398.8	5.0	722	6	CD641598	CD641598 AGENCOURT

## ALIGNMENTS

RESULT 1  
CK284786  
LOCUS EST747508 Nicotiana benthamiana mixed tissue cDNA library, linear EST 02-AUG-2004  
DEFINITION normalised, full-length Nicotiana benthamiana cDNA clone NEMAQ41 5', end, mRNA sequence.  
ACCESSION CK284786  
VERSION CK284786.1 GI:39858698  
KEYWORDS EST.  
SOURCE Nicotiana benthamiana  
ORGANISM Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Nicotiana.  
REFERENCE 1 (bases 1 to 935)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Skaskawicz, B., Jin, H. and Baker, B.  
TITLE Generation of EST sequences from Nicotiana benthamiana  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST747509  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

1..935  
Location/Qualifiers  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NEMAQ41"  
/tissue\_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from



these tissues and pooled in approximately equal molar amounts."

ORIGIN	Query Match	10.0%; Score 795.4; DB 7; Length 935;	
	Best Local Similarity 99.9%; Pred. No. 3.6e-174;		
	Matches 796; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	386	CATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGCCCGCTTGGGTGGAGAGGCTATT	445
Db	31	CATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGCCCGCTTGGGTGGAGAGGCTATT	90
Qy	446	CGGCTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCCGCCGTGTTCCGGGTGTC	505
Db	91	CGGCTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCCGCCGTGTTCCGGGTGTC	150
Qy	506	AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACGACCTTCGGTGCCTCGAATGAACCT	565
Db	151	AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACGACCTTCGGTGCCTCGAATGAACCT	210
Qy	566	GCAGGACGAGCAGCGCGGTATCTGGCTGGCCACGACGGGGGTTCTTTGGCGAGCTGT	625
Db	211	GCAGGACGAGCAGCGCGGTATCTGGCTGGCCACGACGGGGGTTCTTTGGCGAGCTGT	270
Qy	626	GCTCGAGCTTGTCACTGAACGGGGAAGGACTGGCTGCTATTGGCGGAAGTGCCGGGGCA	685
Db	271	GCTCGAGCTTGTCACTGAACGGGGAAGGACTGGCTGCTATTGGCGGAAGTGCCGGGGCA	330
Qy	686	GGATCTCTCTCATCTCACCTTGCTCTCGCGAGAAGTAGTCCATCATGCTCATGCAAT	745
Db	331	GGATCTCTCTCATCTCACCTTGCTCTCGCGAGAAGTAGTCCATCATGCTCATGCAAT	390
Qy	746	CGCGCGCTGCATAACGCTTGATCGGCTACCTGCCATTCGACCAACGCAAAACATCG	805
Db	391	CGCGCGCTGCATAACGCTTGATCGGCTACCTGCCATTCGACCAACGCAAAACATCG	450
Qy	806	CATCGACGAGCAGCTACTCGGATGGNAGCGGCTCTTGTCATCAGGATGATCTGACGGA	865
Db	451	CATCGACGAGCAGCTACTCGGATGGNAGCGGCTCTTGTCATCAGGATGATCTGACGGA	510
Qy	866	AGAGCATCAGGGGCTCGCGCAGCCGAACCTGTTCCCGCAGGCTCAAGCGCGCATGCCCGA	925
Db	511	AGAGCATCAGGGGCTCGCGCAGCCGAACCTGTTCCCGCAGGCTCAAGCGCGCATGCCCGA	570
Qy	926	CGCGCAGGATCTGTGTGTGACCCATGCGGATGCTGTCTCCGAAATATCATGTGTGAAAA	985
Db	571	CGCGCAGGATCTGTGTGTGACCCATGCGGATGCTGTCTCCGAAATATCATGTGTGAAAA	630
Qy	986	TGGCCGCTTTCTGGATTTCATCGACTGTGCCGCTGGGTGGCGAACCGCTATCAGGA	1045
Db	631	TGGCCGCTTTCTGGATTTCATCGACTGTGCCGCTGGGTGGCGAACCGCTATCAGGA	690
Qy	1046	CATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTTGGCGGGAATGGGCTCACCGCTT	1105
Db	691	CATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTTGGCGGGAATGGGCTCACCGCTT	750
Qy	1106	CTCTGCTCTTTACGGTATCGCGCTCCCGATTCGACGGCATCGCCTTCATCGCCTTCT	1165
Db	751	CTCTGCTCTTTACGGTATCGCGCTCCCGATTCGACGGCATCGCCTTCATCGCCTTCT	810
Qy	1166	TGACGAGTTCTTCTGAG 1:182	
Db	811	TGACGAGTTCTTCTGAG 827	

RESULT 2	CK291519	804 bp	linear	EST 02-AUG-2004
LOCUS	EST754233	Nicotiana benthamiana mixed tissue	cdna library	
DEFINITION	normalized, full-length Nicotiana benthamiana cDNA clone NEMC276 5'			
ACCESSION	CK291519	end, mRNA sequence.		
VERSION	CK291519.1	GI:39872047		

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1. J. H. ...	...
2. J. H. ...	...
3. J. H. ...	...
4. J. H. ...	...
5. J. H. ...	...
6. J. H. ...	...
7. J. H. ...	...
8. J. H. ...	...
9. J. H. ...	...
10. J. H. ...	...
11. J. H. ...	...
12. J. H. ...	...
13. J. H. ...	...
14. J. H. ...	...
15. J. H. ...	...
16. J. H. ...	...
17. J. H. ...	...
18. J. H. ...	...
19. J. H. ...	...
20. J. H. ...	...
21. J. H. ...	...
22. J. H. ...	...
23. J. H. ...	...
24. J. H. ...	...
25. J. H. ...	...
26. J. H. ...	...
27. J. H. ...	...
28. J. H. ...	...
29. J. H. ...	...
30. J. H. ...	...
31. J. H. ...	...
32. J. H. ...	...
33. J. H. ...	...
34. J. H. ...	...
35. J. H. ...	...
36. J. H. ...	...
37. J. H. ...	...
38. J. H. ...	...
39. J. H. ...	...
40. J. H. ...	...
41. J. H. ...	...
42. J. H. ...	...
43. J. H. ...	...
44. J. H. ...	...
45. J. H. ...	...
46. J. H. ...	...
47. J. H. ...	...
48. J. H. ...	...
49. J. H. ...	...
50. J. H. ...	...
51. J. H. ...	...
52. J. H. ...	...
53. J. H. ...	...
54. J. H. ...	...
55. J. H. ...	...
56. J. H. ...	...
57. J. H. ...	...
58. J. H. ...	...
59. J. H. ...	...
60. J. H. ...	...
61. J. H. ...	...
62. J. H. ...	...
63. J. H. ...	...
64. J. H. ...	...
65. J. H. ...	...
66. J. H. ...	...
67. J. H. ...	...
68. J. H. ...	...
69. J. H. ...	...
70. J. H. ...	...
71. J. H. ...	...
72. J. H. ...	...
73. J. H. ...	...
74. J. H. ...	...
75. J. H. ...	...
76. J. H. ...	...
77. J. H. ...	...
78. J. H. ...	...
79. J. H. ...	...
80. J. H. ...	...
81. J. H. ...	...
82. J. H. ...	...
83. J. H. ...	...
84. J. H. ...	...
85. J. H. ...	...
86. J. H. ...	...
87. J. H. ...	...
88. J. H. ...	...
89. J. H. ...	...
90. J. H. ...	...
91. J. H. ...	...
92. J. H. ...	...
93. J. H. ...	...
94. J. H. ...	...
95. J. H. ...	...
96. J. H. ...	...
97. J. H. ...	...
98. J. H. ...	...
99. J. H. ...	...
100. J. H. ...	...

FEATURES	source
----------	--------

EST.  
Nicotiana benthamiana  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanaceae; Nicotiana.  
1 (bases 1 to 804)  
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,  
Staekowicz,B., Jin,H. and Baker,B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/.  
Seq primer: ATT TAG GTG ACA CTA TAG.

## ORIGIN

	Query Match	9.9%;	Score 794;	DB 7;	Length 804;
	Best Local Similarity	100.0%;	Pred. No. 7.5e-174;		
	Matches 794;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	389	GATTGAACAAGATGGATTGCACGACGAGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGG	448		
Db	2	GATTGAACAAGATGGATTGCACGAGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGG	61		
Qy	449	CTATGACTGGGCAACAACAGACAATCGGTGCTCTGATCCGCGGTGTTCGGGCTGTCAAGC	508		
Db	62	CTATGACTGGGCAACAACAGACAATCGGTGCTCTGATCCGCGGTGTTCGGGCTGTCAAGC	121		
Qy	509	GCAGGGGGCCCGGTTCTTTTGTGCAAGACCGACTGCTCCGGTGCCCTGAATGAACATGCA	568		
Db	122	GCAGGGGGCCCGGTTCTTTTGTGCAAGACCGACTGCTCCGGTGCCCTGAATGAACATGCA	181		
Qy	569	GGACGAGCGACGCGGGCTATCGTGGCTGGCCACGACGGCGGTTCCTTCGGCAGCTGTGCT	628		
Db	182	GGACGAGCGACGCGGGCTATCGTGGCTGGCCACGACGGCGGTTCCTTCGGCAGCTGTGCT	241		
Qy	629	CGAGTTGTCACTGAAGGGGAAGGACTGCGTGTATTGGGCGAAGTGCCTGGGGCAGGA	688		
Db	242	CGAGTTGTCACTGAAGCGGAAGGACTGCGTGTATTGGGCGAAGTGCCTGGGGCAGGA	301		
Qy	689	TCTCCTGTCACTCACTTGTCTCTGCGCGAGAAAGTATCCATCATCGGTGATGCATATGGC	748		
Db	302	TCTCCTGTCACTCACTTGTCTCTGCGCGAGAAAGTATCCATCATCGGTGATGCATATGGC	361		
Qy	749	CGCGCTGCATACGTTGATCCGGGTACTCTGCCCATTCGACCAACCAAGCGAAACATTCGCAT	808		
Db	362	CGCGCTGCATACGTTGATCCGGGTACTCTGCCCATTCGACCAACCAAGCGAAACATTCGCAT	421		

QY 809 CGAGGACAGTACTCGGATGGAGCCGGTCTTCTCGATCAGGATGATCTGGACGAGA 868  
 |||||  
 Db 422 CGAGGACAGTACTCGGATGGAGCCGGTCTTCTCGATCAGGATGATCTGGACGAGA 481  
 |||||  
 QY 869 GCATCAGGGGCTCGCGCCAGCCGAACTGTTGCCAGGCTCAAGGCGCGCATGCCGACGG 928  
 |||||  
 Db 482 GCATCAGGGGCTCGCGCCAGCCGAACTGTTGCCAGGCTCAAGGCGCGCATGCCGACGG 541  
 |||||  
 QY 929 CGAGGATCTCTGTGACCCATGGGATGCTGCTTGGCCGAATATCATGGTGGAAAAATGG 988  
 |||||  
 Db 542 CGAGGATCTCTGTGACCCATGGGATGCTGCTTGGCCGAATATCATGGTGGAAAAATGG 601  
 |||||  
 QY 989 CCGCTTTTCTGATTCATCGACTGTGGCCGCTGGTGTGGGACCCGCTATCAGGACAT 1048  
 |||||  
 Db 602 CCGCTTTTCTGATTCATCGACTGTGGCCGCTGGTGTGGGACCCGCTATCAGGACAT 661  
 |||||  
 QY 1049 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCT 1108  
 |||||  
 Db 662 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCT 721  
 |||||  
 QY 1109 GGTGCTTTACGGTATCGCGCTCCGATTCGAGCGGATCGCCTTCTATGCGCTTCTTGA 1168  
 |||||  
 Db 722 GGTGCTTTACGGTATCGCGCTCCGATTCGAGCGGATCGCCTTCTATGCGCTTCTTGA 781  
 |||||  
 QY 1169 CGAGTCTTCTTGAG 1182  
 |||||  
 Db 782 CGAGTCTTCTTGAG 795  
 |||||

## RESULT 3

CK287297

## LOCUS

DEFINITION CK287297 856 bp mRNA linear EST 02-AUG-2004  
 EST750019 Nicotiana benthamiana mixed tissue cDNA library,  
 normalized, full-length Nicotiana benthamiana cDNA clone NEMB815 5'  
 end, mRNA sequence.

## ACCESSION

CK287297

## VERSION

CK287297.1

## KEYWORDS

GI:39863696

## SOURCE

EST.

## ORGANISM

Nicotiana benthamiana

## REFERENCE

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
 Staskiewicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST750020

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>

Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

source

1..856

Location/Qualifiers

/organism="Nicotiana benthamiana"

/mol\_type="mRNA"

/db\_xref="taxon:4100"

/clone="NEMB815"

/tissue\_type="abiotic and biotic stress-treated leaves,  
 callus tissue and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
 library, normalized, full-length"/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: RNA was isolated from Nicotiana benthamiana  
 tissues that include callus, roots from liquid culture  
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen

challenged leaves (*Pseudomonas syringae* pv *tomato* 12 hr;  
*Xanthomonas campestris* pv *campestris* 12 hr, 18hr;  
*Pseudomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas*  
*campestris* pv *vesicatoria* 18hr). RNA was isolated from  
 these tissues and pooled in approximately equal molar  
 amounts."

## ORIGIN

Query Match 9.9%; Score 794; DB 7; Length 856;

Best Local Similarity 100.0%; Pred. No. 7.6e-174;

Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 389 GATTGAACAAGATGGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGCTATTTCGG 448

Db 14 GATTGAACAAGATGGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGCTATTTCGG 73

QY 449 CTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTCAAG 508

Db 74 CTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTCAAG 133

QY 509 GCAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAAGTCA 568

Db 134 GCAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAAGTCA 193

QY 569 GGACGAGGACGCGCGCTATCGTGGCTGGCCACAGCGGCGTTCCTTGGCGAGCTGTGCT 628

Db 194 GGACGAGGACGCGCGCTATCGTGGCTGGCCACAGCGGCGTTCCTTGGCGAGCTGTGCT 253

QY 629 CGAGCTTGTCTCACTGAAGCGGGAAGGAGTGTCTCTATTTGGCGGAAGTGTCCGGGCG 688

Db 254 CGAGCTTGTCTCACTGAAGCGGGAAGGAGTGTCTCTATTTGGCGGAAGTGTCCGGGCG 313

QY 689 TCTCTGTCTATCTCACCTTGTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCG 748

Db 314 TCTCTGTCTATCTCACCTTGTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCG 373

QY 749 CGGCTGCATACGCTTGTATCCGGCTACTGCGGCTCAAGCGCGCATGCCCCGACGG 808

Db 374 CGGCTGCATACGCTTGTATCCGGCTACTGCGGCTCAAGCGCGCATGCCCCGACGG 433

QY 809 CGAGGACGAGTACTCGGATGGAGCCGGTCTTGTGATCAGATGATCTGGACGAGA 868

Db 434 CGAGGACGAGTACTCGGATGGAGCCGGTCTTGTGATCAGATGATCTGGACGAGA 493

QY 869 GCATCAGGGGCTCGCGCCAGCCGAACTGTTGCCAGGCTCAAGCGCGCATGCCCCGACGG 928

Db 494 GCATCAGGGGCTCGCGCCAGCCGAACTGTTGCCAGGCTCAAGCGCGCATGCCCCGACGG 553

QY 929 CGAGGATCTCTGTGTGACCCATGGCGCATGCTTGTCCGAATATCATGGTGGAAAAATGG 988

Db 554 CGAGGATCTCTGTGTGACCCATGGCGCATGCTTGTCCGAATATCATGGTGGAAAAATGG 613

QY 989 CCGCTTTTCTGGATTCATCCACTGTGCCCGCTGGGTGGCGACCGCTATCAGGACAT 1048

Db 614 CCGCTTTTCTGGATTCATCCACTGTGCCCGCTGGGTGGCGACCGCTATCAGGACAT 673

QY 1049 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCT 1108

Db 674 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCT 733

QY 1109 CGTGTCTTACGGTATCGCCGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 1168

Db 734 CGTGTCTTACGGTATCGCCGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 793

QY 1169 CGAGTCTTCTTGAG 1182

Db 794 CGAGTCTTCTTGAG 807

RESULT 4

CK287930

LOCUS

DEFINITION CK287930 910 bp mRNA linear EST 02-AUG-2004

EST750652 Nicotiana benthamiana mixed tissue cDNA library,

normalized, full-length Nicotiana benthamiana cDNA clone NBMC75 5' end, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CK287930  
CK287930.1 GI:39864940  
EST.

Nicotiana benthamiana  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE  
AUTHORS

1 (bases 1 to 910)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Staskawicz, B., Jin, H. and Baker, B.

TITLE  
JOURNAL

Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)

COMMENT

Other ESTs: EST750653  
Contact: Robin Buell

The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

Location/Qualifiers  
1..910

/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMC75"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

ORIGIN

Query Match 9.9%; Score 794; DB 7; Length 910;  
Best Local Similarity 100.0%; Pred. No. 7.7e-174;  
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

389 GATTGAACAGATGATGACGAGATTTCTCGGCGCTTGGTGAGAGGCTATTCGG 448

4 GATTGAACAGATGATGACGAGATTTCTCGGCGCTTGGTGAGAGGCTATTCGG 63

449 CTATGACTGGGCAACAGACAATCGGCTGCTGTGATGCCCGCTTTCGGCTGTCAGC 508

64 CTATGACTGGGCAACAGACAATCGGCTGCTGTGATGCCCGCTTTCGGCTGTCAGC 123

509 GCAGGGGCGCGCGGTTCTTTTGTGCAAGACGACCTGTCCGGTGCCCTGAATGAACCTGCA 568

124 GCAGGGGCGCGCGGTTCTTTTGTGCAAGACGACCTGTCCGGTGCCCTGAATGAACCTGCA 183

569 GGACGAGGACGCGGCTATCGTGGTGGGACAGCGGGCTTCTTGGCAGCTGTGCT 628

184 GGACGAGGACGCGGCTATCGTGGTGGGACAGCGGGCTTCTTGGCAGCTGTGCT 243

629 CGACGTTGTCACTGAACGGGAAGGAGTGTGCTCTATTGGGCGAAGTCCGGGGGACGGA 688

244 CGACGTTGTCACTGAACGGGAAGGAGTGTGCTCTATTGGGCGAAGTCCGGGGGACGGA 303

689 TCTCTCTCATCTCACTTGTCTCTCGCGAGAAAGTATCCATCATGCTGATGAATGCG 748

304 TCTCTCTCATCTCACTTGTCTCTCGCGAGAAAGTATCCATCATGCTGATGAATGCG 363

749 GCGGCTGCATACGCTTGTATCGGCTACCTGCCATTTCGACCAACGAAACATCGCAT 808

364 GCGGCTGCATACGCTTGTATCGGCTACCTGCCATTTCGACCAACGAAACATCGCAT 423

809 CGAGCGAGCAGCTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGA 868

424 CGAGCGAGCAGCTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGA 483

869 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGCGCGCATGCCGACCG 928

484 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGCGCGCATGCCGACCG 543

929 CGAGGATCTCGTCTGACCCATGCGCATGCTGCTTCCGAAATATCATGCTGGAATG 988

544 CGAGGATCTCGTCTGACCCATGCGCATGCTGCTTCCGAAATATCATGCTGGAATG 603

989 CCGCTTTTCTGGATTCATCGATGCGCGCTGGTGTGGCGGACCGCTATCAGGACAT 1048

604 CCGCTTTTCTGGATTCATCGATGCGCGCTGGTGTGGCGGACCGCTATCAGGACAT 663

1049 AGCGTTGGCTACCGCTGATATTCGCAAGAGCTTGGCGGAGTGGCTGACCGCTTCCT 1108

664 AGCGTTGGCTACCGCTGATATTCGCAAGAGCTTGGCGGAGTGGCTGACCGCTTCCT 723

1109 CGTGTCTTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 1168

724 CGTGTCTTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 783

1169 CGAGTTCTTCTGAG 1182

784 CGAGTTCTTCTGAG 797

RESULT 5

LOCUS

DEFINITION

CK291799 933 bp mRNA linear EST 02-AUG-2004  
EST754513 Nicotiana benthamiana mixed tissue cDNA library,  
normalized, full-length Nicotiana benthamiana cDNA clone NBMC477 5'  
end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS

OTHER ESTs: EST754514

Unpublished (2003)

Generation of EST sequences from Nicotiana benthamiana

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..933

/organism="Nicotiana benthamiana"

/mol\_type="mRNA"

/db\_xref="taxon:4100"

/clone="NBMC477"

/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

```
ORIGIN
Query Match          9.9%; Score 794; DB 7; Length 933;
Best Local Similarity 100.0%; Pred. No. 7.7e-174;
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 389 GATTGAACAAGATGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCG 448
DB 46 GATTGAACAAGATGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCG 105
QY 449 CTATGACTGGGACACAGACAAATCGGTCTCTGATCGCGCGGTTCGGGTGCTGACG 508
DB 106 CTATGACTGGGACACAGACAAATCGGTCTCTGATCGCGCGGTTCGGGTGCTGACG 165
QY 509 GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAACTGCA 568
DB 166 GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAACTGCA 225
QY 569 GGAAGGACGCGGGTATCGTGGTGGCCACGACGCGGCTTCTTCCGCGAGCTGTGCT 628
DB 226 GGAAGGACGCGGGTATCGTGGTGGCCACGACGCGGCTTCTTCCGCGAGCTGTGCT 285
QY 629 CGAGTTGTCTACTGAAGCGGGAAGGACTGCTGCTATGCGGGAAGTCCGGGCGAGGA 688
DB 286 CGAGTTGTCTACTGAAGCGGGAAGGACTGCTGCTATGCGGGAAGTCCGGGCGAGGA 345
QY 689 TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATCATGGTGTGATGAATGCG 748
DB 346 TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATCATGGTGTGATGAATGCG 405
QY 749 GCGGCTGCATACGCTTGTATCGGCTTACCTGCCATTCGACCACCAAGCGAAACATCGCAT 808
DB 406 GCGGCTGCATACGCTTGTATCGGCTTACCTGCCATTCGACCACCAAGCGAAACATCGCAT 465
QY 809 CGAGCGAGCAGTACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGGAAGA 868
DB 466 CGAGCGAGCAGTACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGGAAGA 525
QY 869 GCATCAGGGGCTCGCGCCAGCGCAACTGTTCCGCGAGGCTCAAGCGCGCATGCCGACGG 928
DB 526 GCATCAGGGGCTCGCGCCAGCGCAACTGTTCCGCGAGGCTCAAGCGCGCATGCCGACGG 585
QY 929 CGAGGATCTCGTGTGATCCCATGCGGATGCTGCTTCCGCAATATCATGTGGAAATGG 988
DB 586 CGAGGATCTCGTGTGATCCCATGCGGATGCTGCTTCCGCAATATCATGTGGAAATGG 645
QY 989 CCGGTTTTCTGATTCATGCACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACAT 1048
DB 646 CCGGTTTTCTGATTCATGCACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACAT 705
QY 1049 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCT 1108
DB 706 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCT 765
QY 1109 CGTGCTTTACGGTATCGCGCTCCCGATTCGACGCGCATGCGCTTCTATGCGCTTCTTGA 1168
DB 766 CGTGCTTTACGGTATCGCGCTCCCGATTCGACGCGCATGCGCTTCTATGCGCTTCTTGA 825
QY 1169 CGAGTTCTTCTGAG 1182
DB 826 CGAGTTCTTCTGAG 839
```

## RESULT 6

CK256977

LOCUS

DEFINITION

CK256977 936 bp mRNA linear EST 30-JUL-2004  
EST740614 potato callus cDNA library, normalized and full-length  
Solanum tuberosum cDNA clone POC170 5' end, mRNA sequence.

ACCESSION

CK256977.1 GI:39813957

VERSION

EST.

SOURCE

Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@igr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

Location/Qualifiers

1..936

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clones="POCD170"

/tissue\_type="callus"

/lab\_host="DH10B-Tona"

/clone\_lib="potato callus cDNA library, normalized and

full-length"

/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Solanum tuberosum var.

Kennebec callus tissue grown on solid media."

ORIGIN

Query Match

Best Local Similarity

Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 389 GATTGAACAAGATGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCG 448

DB 14 GATTGAACAAGATGATTGACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCG 73

QY 449 CTATGACTGGGACACAGACAAATCGGCTGCTGTGATGCGCGCGCTTCCGGCTGTGCTCAGC 508

DB 74 CTATGACTGGGACACAGACAAATCGGCTGCTGTGATGCGCGCGCTTCCGGCTGTGCTCAGC 133

QY 509 GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTTCGCGAGCTGTGCT 628

DB 134 GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTTCGCGAGCTGTGCT 193

QY 569 GGAAGGACGCGCGGCTATCGTGGCTGGCCACAGACGGGCGTTCCTTTCGCGAGCTGTGCT 628

DB 194 GGAAGGACGCGCGGCTATCGTGGCTGGCCACAGACGGGCGTTCCTTTCGCGAGCTGTGCT 253

QY 629 GGAAGGACGCGCGGCTATCGTGGCTGGCCACAGACGGGCGTTCCTTTCGCGAGCTGTGCT 688

DB 254 GGAAGGACGCGCGGCTATCGTGGCTGGCCACAGACGGGCGTTCCTTTCGCGAGCTGTGCT 313

QY 689 TCTCCCTGCTATCTCACCTTGTCTCCCGGAGAAAGTATCCATCATGCTGATGCAATGCG 748

DB 314 TCTCCCTGCTATCTCACCTTGTCTCCCGGAGAAAGTATCCATCATGCTGATGCAATGCG 373

QY 749 GCGGCTGCATACGCTTGTATCCGGCTACCTGCCCATTCGACCAAGCGAAACATCGCAT 808

DB 374 GCGGCTGCATACGCTTGTATCCGGCTACCTGCCCATTCGACCAAGCGAAACATCGCAT 433

QY 809 CGAGGAGCAGCTACTCGATGGAAGCCGGTCTTGTGATCAGGATGATCTGACGAGA 868  
|||||  
Db 434 CGAGGAGCAGCTACTCGATGGAAGCCGGTCTTGTGATCAGGATGATCTGACGAGA 493  
|||||  
QY 869 GCATCAGGGGCTCGGCCAGCCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCGACGG 928  
|||||  
Db 494 GCATCAGGGGCTCGGCCAGCCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCGACGG 553  
|||||  
QY 929 CGAGGATCTGCTGTCGACCCATGCGGATGCGCTGCTTCCGATATATCATGTTGGAATGG 988  
|||||  
Db 554 CGAGGATCTGCTGTCGACCCATGCGGATGCGCTGCTTCCGATATATCATGTTGGAATGG 613  
|||||  
QY 989 CGCTTTTCTGATTCATGACTGTCGCGGCTGGGTGGGACCGCTATCAGGACAT 1048  
|||||  
Db 614 CGCTTTTCTGATTCATGACTGTCGCGGCTGGGTGGGACCGCTATCAGGACAT 673  
|||||  
QY 1049 AGCGTTGGTACCGCTGATATTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGTTCT 1108  
|||||  
Db 674 AGCGTTGGTACCGCTGATATTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGTTCT 733  
|||||  
QY 1109 CGTGTCTTACGGTATCGCGCTCCGATTCGAGCGCATGCGCTTCTATGCGCTTCTTGA 1168  
|||||  
Db 734 CGTGTCTTACGGTATCGCGCTCCGATTCGAGCGCATGCGCTTCTATGCGCTTCTTGA 793  
|||||  
QY 1169 CGAGTCTTCTGAG 1182  
|||||  
Db 794 CGAGTCTTCTGAG 807  
|||||

RESULT 7  
CK298208  
LOCUS  
DEFINITION  
EST760922 Nicotiana benthamiana mixed tissue cDNA library  
normalized, full-length Nicotiana benthamiana cDNA clone NEMDE30 5'  
end, mRNA sequence.

ACCESSION  
CK298208  
VERSION  
CK298208.1 GI:39885354  
KEYWORDS  
EST.

SOURCE  
Nicotiana benthamiana  
ORGANISM  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE  
AUTHORS  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Staskiewicz, B., Jin, H. and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Other ESTs: EST760923  
Contact: Robin Buell

TITLE  
JOURNAL  
COMMENT  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
source  
1..947  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NEMDE30"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/note="vector: pCWVSp6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;

Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

## ORIGIN

Query Match 9.9%; Score 794; DB 7; Length 947;  
Best Local Similarity 100.0%; Pred. No. 7.7e-174;  
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 389 GATTGAACAAGATGATTCACACGAGTTCTCCGCCGCTTGGGTGAGAGGCTATTCGG 448  
|||||  
Db 17 GATTGAACAAGATGATTCACACGAGTTCTCCGCCGCTTGGGTGAGAGGCTATTCGG 76  
|||||  
QY 449 CTATGACTGGGCACACACACATCGGCTCTGATGCCGCTGCTCCGGCTGTCAGC 508  
|||||  
Db 77 CTATGACTGGGCACACACACATCGGCTCTGATGCCGCTGCTCCGGCTGTCAGC 136  
|||||  
QY 509 CGAGGGGCGCCGGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAATGAAC 568  
|||||  
Db 137 CGAGGGGCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAATGAAC 196  
|||||  
QY 569 GGAAGGAGGCGCGCTATCGTGGTGGCAGCAGCGGGCTTCCTTGGCAGCTGTGCT 628  
|||||  
Db 197 GGAAGGAGGCGCGCTATCGTGGTGGCAGCAGCGGGCTTCCTTGGCAGCTGTGCT 256  
|||||  
QY 629 CGACGTTGTCACTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGCCGGGCGAGGA 688  
|||||  
Db 257 CGACGTTGTCACTGAAGCGGGAAGGACCTGGCTGCTATTGGGCGAAGTGCCGGGCGAGGA 316  
|||||  
QY 689 TCTCTGTCTCATCTCACTTGTCTCTCCGAGAAAGTATCCATCATGGTGTATCAATGCG 748  
|||||  
Db 317 TCTCTGTCTCATCTCACTTGTCTCTCCGAGAAAGTATCCATCATGGTGTATCAATGCG 376  
|||||  
QY 749 GCGGCTGCATACGCTTGTATCCGGCTACCTGCCCATTTCGACCAACGCGAAATCGCAT 808  
|||||  
Db 377 GCGGCTGCATACGCTTGTATCCGGCTACCTGCCCATTTCGACCAACGCGAAATCGCAT 436  
|||||  
QY 809 CGAGCGAGCAGCTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAGA 868  
|||||  
Db 437 CGAGCGAGCAGCTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAGA 496  
|||||  
QY 869 GCATCAGGGGCTCGGCCAGCCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCGACGG 928  
|||||  
Db 497 GCATCAGGGGCTCGGCCAGCCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCGACGG 556  
|||||  
QY 929 CGAGGATCTGCTGTCGACCCATGCGGCTGCTTCCGGAATATCATGTTGGAATGG 988  
|||||  
Db 557 CGAGGATCTGCTGTCGACCCATGCGGCTGCTTCCGGAATATCATGTTGGAATGG 616  
|||||  
QY 989 CGGCTTTCTGATTCATCGACTGTGCGCGTGGGTGGCGACCGCTATCAGGACAT 1048  
|||||  
Db 617 CGGCTTTCTGATTCATCGACTGTGCGCGTGGGTGGCGACCGCTATCAGGACAT 676  
|||||  
QY 1049 AGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCT 1108  
|||||  
Db 677 AGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCT 736  
|||||  
QY 1109 CGTGTCTTACGGTATCGCGCTCCGATTCGAGCGCATGCGCTTCTATGCGCTTCTTGA 1168  
|||||  
Db 737 CGTGTCTTACGGTATCGCGCTCCGATTCGAGCGCATGCGCTTCTATGCGCTTCTTGA 796  
|||||  
QY 1169 CGAGTCTTCTGAG 1182  
|||||  
Db 797 CGAGTCTTCTGAG 810  
|||||

## RESULT 8

CK283361

LOCUS

DEFINITION

EST746083 Nicotiana benthamiana mixed tissue cDNA library,  
normalized, full-length Nicotiana benthamiana cDNA clone NEMAG50 5'

CK283361 954 bp mRNA linear EST 02-AUG-2004  
EST746083 Nicotiana benthamiana mixed tissue cDNA library,  
normalized, full-length Nicotiana benthamiana cDNA clone NEMAG50 5'

```

end, mRNA sequence.
ACCESSION CK283361
VERSION CK283361.1 GI:39855898
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE asterids; lamids; Solanales; Solanaceae; Nicotiana.
JOURNAL 1 (bases 1 to 954)
COMMENT Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
Staskawicz,B., Jin,H. and Baker,B.
Generation of EST sequences from Nicotiana benthamiana
Other ESTs: EST746084
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..954
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMAG50"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana"
ORIGIN
Query Match 9.9%; Score 794; DB 7; Length 954;
Best Local Similarity 100.0%; Pred. No. 7.7e-174;
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 389 GATTGAACAGATGATGACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTGG 448
Db 14 GATTGAACAGATGATGACGAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTGG 73
QY 449 CTATGACTGGGCACACAGACAATCGGCTGCTCTGATGCCCGCTTTCGGCTGTCAGC 508
Db 74 CTATGACTGGGCACACAGACAATCGGCTGCTCTGATGCCCGCTTTCGGCTGTCAGC 133
QY 509 CGAGGGGGCCCGCTTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGATGAAGTCA 568
Db 134 CGAGGGGGCCCGCTTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGATGAAGTCA 193
QY 569 GGACGAGGACGCGGCTATCGTGGTGGCCACAGCGGCGTTCCTTGGCAGCTGTGCT 628
Db 194 GGACGAGGACGCGGCTATCGTGGTGGCCACAGCGGCGTTCCTTGGCAGCTGTGCT 253
QY 629 CGAGCTTGTCTACTAAGCGGGAAGGACTGCTGCTATTGGCGGAAGTCCGGGGCAGGA 688
Db 254 CGAGCTTGTCTACTAAGCGGGAAGGACTGCTGCTATTGGCGGAAGTCCGGGGCAGGA 313
QY 689 TCTCCTGTCTATCTACCTTGTCTCTCGCGGAGAAAGTATCCATCATGCTGATGAATGCG 748
Db 314 TCTCCTGTCTATCTACCTTGTCTCTCGCGGAGAAAGTATCCATCATGCTGATGAATGCG 373

```

```

QY 749 GCGCTGCATACGCTTGATTCGGCTACCTGCCATTTCGACCAACGAAGCAACATCGCAT 808
Db 374 GCGCTGCATACGCTTGATTCGGCTACCTGCCATTTCGACCAACGAAGCAACATCGCAT 433
QY 809 CGAGGACACGCTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTCGACGAAGA 868
Db 434 CGAGGACACGCTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTCGACGAAGA 493
QY 869 GCATCAGGGGTCGCGCCAGCCGAACCTGTCGCCAGGCTCAAGGCGCGCATGCCGACGG 928
Db 494 GCATCAGGGGTCGCGCCAGCCGAACCTGTCGCCAGGCTCAAGGCGCGCATGCCGACGG 553
QY 929 CGAGGATCTCGTGTGAGCCCATGCGGATGCGCTTTCGCCGAATATCATGTGGAAATGG 988
Db 554 CGAGGATCTCGTGTGAGCCCATGCGGATGCGCTTTCGCCGAATATCATGTGGAAATGG 613
QY 989 CCGCTTTTCTCGATTTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 1048
Db 614 CCGCTTTTCTCGATTTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACAT 673
QY 1049 AGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTTGGCGCGGAATGGGCTTCCCT 1108
Db 674 AGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTTGGCGCGGAATGGGCTTCCCT 733
QY 1109 CGTCTTTTACGGTATCGCGCTCCGATTCGAGGATTCGAGGCGATCGCCTTCTTCTTGA 1168
Db 734 CGTCTTTTACGGTATCGCGCTCCGATTCGAGGATTCGAGGCGATCGCCTTCTTCTTGA 793
QY 1169 CGAGTTCTTCTGAG 1182
Db 794 CGAGTTCTTCTGAG 807

```

```

RESULT 9
LOCUS CK288185
DEFINITION EST750907 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NEMBE61 5',
end, mRNA sequence.
ACCESSION CK288185
VERSION CK288185.1 GI:39865462
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 811)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
Staskawicz,B., Jin,H. and Baker,B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..811
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMBE61"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana"

```

tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN

Query Match	9.7%;	Score 778.8;	DB 7;	Length 811;
Best Local Similarity	99.7%;	Pred. No. 2.6e-170;	Indels 0;	Gaps 0;
Matches 780;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

  

QY	386	CATGATTGAACAGATGGATTGCACGAGTTCTCCGGCGCTGGTGGAGAGGCTATT	445
Db	30	CATGATTGAACAGATGGATTGCACGAGTTCTCCGGCGCTGGTGGAGAGGCTATT	89
QY	446	CGGCTATGACTGGGCAACAGCAATCGGCTGCTGATGCGCGGTGTCGGGCTGC	505
Db	90	CGGCTATGACTGGGCAACAGCAATCGGCTGCTGATGCGCGGTGTCGGGCTGC	149
QY	506	AGCGCAGGGGCGCGCGTCTTTTGTCAAGACCGACCTGTCGGGTGCCCTGAATGAAC	565
Db	150	AGCGCAGGGGCGCGCGTCTTTTGTCAAGACCGACCTGTCGGGTGCCCTGAATGAAC	209
QY	566	GCAGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGGCGTTCCTTGCAGCTGT	625
Db	210	GCAGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGGCGTTCCTTGCAGCTGT	269
QY	626	GCTCGAGCTTGTCACTGAACGGGAGGAGTGGCTGCTATTTGGGCGAGTGGCGGCA	685
Db	270	GCTCGAGCTTGTCACTGAACGGGAGGAGTGGCTGCTATTTGGGCGAGTGGCGGCA	329
QY	686	GGATCTCTGCTATCTACCTTGTCTCGCGGAGAAAGTATCCATCATGCTGATGCAAT	745
Db	330	GGATCTCTGCTATCTACCTTGTCTCGCGGAGAAAGTATCCATCATGCTGATGCAAT	389
QY	746	GGCGGGGCTGATACGCTTGCATCGGCTACCTGCCATTCGACACCAAGCAAAATCG	805
Db	390	GGCGGGGCTGATACGCTTGCATCGGCTACCTGCCATTCGACACCAAGCAAAATCG	449
QY	806	CATCGAGCGAGCAGTACTCGATGGAGCGGCTTGTGATCAGATGATCTGGACGA	865
Db	450	CATCGAGCGAGCAGTACTCGATGGAGCGGCTTGTGATCAGATGATCTGGACGA	509
QY	866	AGAGCATCAGGGGCTCGCGCAGCCGAATGTTGCGCAGGCTCAAGCGCGCATGCCGA	925
Db	510	AGAGCATCAGGGGCTCGCGCAGCCGAATGTTGCGCAGGCTCAAGCGCGCATGCCGA	569
QY	926	CGGCGAGGATCTCGTGTGACCCATGCGGATGCTGTTGCCGAATATCATGTGGAAAA	985
Db	570	CGGCGAGGATCTCGTGTGACCCATGCGGATGCTGTTGCCGAATATCATGTGGAAAA	629
QY	986	TGGCGGCTTTCTGGATTATCGACTGTCGGCGGCTGGGTGTCGCGACCGCTATCAGGA	1045
Db	630	TGGCGGCTTTCTGGATTATCGACTGTCGGCGGCTGGGTGTCGCGACCGCTATCAGGA	689
QY	1046	CATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTTGGCGGCGAAATGGGCTGACCGCTT	1105
Db	690	CATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTTGGCGGCGAAATGGGCTGACCGCTT	749
QY	1106	CCTCGTCTTTACGGTATCGCGCTCCCGATTCGCGAGCGATCGCCCTTCTATCGCCTTCT	1165
Db	750	CCTCGTCTTTACGGTATCGCGCTCCCGATTCGCGAGCGATCGCCCTTCTATCGCCTTCT	809
QY	1166	TG 1167	
Db	810	TG 811	

RESULT 10

CK288711  
LOCUS  
DEFINITION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

CK288711 878 bp mRNA linear EST 02-AUG-2004  
EST751433 Nicotiana benthamiana mixed tissue cDNA library.  
normalized, full-length Nicotiana benthamiana cDNA clone NBM149 5'  
end, mRNA sequence.  
CK288711  
CK288711.1 GI:39866496  
EST.  
Nicotiana benthamiana  
Nicotiana benthamiana  
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 878)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Taskawicz, B., Jin, H., and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers  
1..878  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBM149"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

ORIGIN

Query Match	7.3%;	Score 587;	DB 7;	Length 878;
Best Local Similarity	100.0%;	Pred. No. 1.5e-125;	Indels 0;	Gaps 0;
Matches 587;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

QY	596	GGCCACGACGCGGGCTTCCTTGCAGCTGTCGAGCTGTCTCACTGAAGCGGAAGGGA	655
Db	1	GGCCACGACGCGGGCTTCCTTGCAGCTGTCGAGCTGTCTCACTGAAGCGGAAGGGA	60
QY	656	CTGGCTGCTATTGGCGGAGTGGCGGGCAGGATCTCTGTCATCTCACCTTGTCTCTGC	715
Db	61	CTGGCTGCTATTGGCGGAGTGGCGGGCAGGATCTCTGTCATCTCACCTTGTCTCTGC	120
QY	716	CGAAGAAAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGGCTAC	775
Db	121	CGAAGAAAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGGCTAC	180
QY	776	CTGCCATTCCGACCAACGCGAAACATCCGATCGAGCGAGCAGTACTCGGATGGAGGC	835
Db	181	CTGCCATTCCGACCAACGCGAAACATCCGATCGAGCGAGCAGTACTCGGATGGAGGC	240
QY	836	CGGTCTTGTGATCAGGATGATCTGGACGAGGAGCATCAGGGGCTCGCCAGCCGAACT	895
Db	241	CGGTCTTGTGATCAGGATGATCTGGACGAGGAGCATCAGGGGCTCGCCAGCCGAACT	300



896 GTTCGCCAGGCTCAAGCGCGCATGCCGACGGGAGGATCTCGTGTGACCCATGGCGA 955  
 |||||  
 Db 301 GTTCGCCAGGCTCAAGCGCGCATGCCGACGGGAGGATCTCGTGTGACCCATGGCGA 360  
 |||||  
 Qy 956 TGCGTGTTCGCCGAATATCATGTGTGGAATAATGGCCGCTTTTCTGGAATCATCGACTGTG 1015  
 Db 361 TGCGTGTTCGCCGAATATCATGTGTGGAATAATGGCCGCTTTTCTGGAATCATCGACTGTG 420  
 |||||  
 Qy 1016 CCGGCTGGGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGATATTGCTGA 1075  
 Db 421 CCGGCTGGGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGATATTGCTGA 480  
 |||||  
 Qy 1076 AGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACCGTATCGCGCTCCCGA 1135  
 Db 481 AGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACCGTATCGCGCTCCCGA 540  
 |||||  
 Qy 1136 TTGCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAG 1182  
 Db 541 TTGCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAG 587  
 |||||

RESULT 11  
 AQ361914  
 LOCUS  
 DEFINITION  
 mgxb0005K01f CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
 clone mgxb0005K01f, genomic survey sequence.  
 ACCESSION  
 AQ361914  
 VERSION  
 AQ361914.1 GI:4211753  
 KEYWORDS  
 GSS.  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

1 (bases 1 to 789)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: TAATACGACTCCTATAGGG  
 Class: BAC ends  
 High quality sequence start: 41  
 High quality sequence stop: 392.  
 Location/Qualifiers

1. .789  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0005K01f"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with  
 a haploid genome (n=7) of approximately 40 Mbp. Rice  
 blast is an important model fungal pathogen for studying  
 numerous aspects of the fungal-host interaction. In  
 order to facilitate genome wide analysis, a BAC library  
 containing 9216 clones with an average insert size of 130  
 kbp was constructed. This library represents greater  
 than 25x genome coverage. High density colony filters  
 are available upon request."

ORIGIN

Query Match 7.2%; Score 577.4; DB 8; Length 789;  
 Best Local Similarity 97.8%; Pred. No. 2.6e-123;  
 Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 Qy 581 GCAGCTATCGTGGCTGCGCACGACGGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTTCAC 640  
 |||||  
 Db 39 GCGCTATCGTGGCTGCGCACGACGGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTTCAC 98  
 |||||  
 Qy 641 TGAAGCGGGGAAGGACCTGGCTGTATTGGGCGAAGTGCOCGGGCGAGATCTCTGTTCATC 700  
 |||||  
 Db 99 TGAAGCGGGGAAGGACCTGGCTGTATTGGGCGAAGTGCOCGGGCGAGATCTCTGTTCATC 158  
 |||||  
 Qy 701 TCACCTTGTCTCCCTGCGAGAAAGTATCCATCATGGCTGATGCAATCGCGGGCTGTGCATAC 760  
 |||||  
 Db 159 TCACCTTGTCTCCCTGCGAGAAAGTATCCATCATGGCTGATGCAATCGCGGGCTGTGCATAC 218  
 |||||  
 Qy 761 GCTTGATCCGGCTACCTGCCCATTCGACCAACCAAGCAATATCGCATCGAGCGAGCAGC 820  
 |||||  
 Db 219 GCTTGATCCGGCTACCTGCCCATTCGACCAACCAAGCAATATCGCATCGAGCGAGCAGC 278  
 |||||  
 Qy 821 TACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGACGAGAGATCAGGGGCT 880  
 |||||  
 Db 279 TACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGACGAGAGATCAGGGGCT 338  
 |||||  
 Qy 881 CCGCGCAGCGAACTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGT 940  
 |||||  
 Db 339 CCGCGCAGCGAACTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGT 398  
 |||||  
 Qy 941 CGTGACCCATGGCGATGCGCTCTTGGCGAATATCATGTTGGAATAATGGCGCTTTCTTG 1000  
 |||||  
 Db 399 CGTGACCCATGGCGATGCGCTCTTGGCGAATATCATGTTGGAATAATGGCGCTTTCTTG 458  
 |||||  
 Qy 1001 ATTTCATGCACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATACGTTGGCTAC 1060  
 |||||  
 Db 459 ATTTCATGCACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATACGTTGGCTAC 518  
 |||||  
 Qy 1061 CCGTGATATGCTGAAGAGCTTGGCGCGAATGGGCTACCGCTTCTCTGCTTTACGG 1120  
 |||||  
 Db 519 CCGTGATATGCTGAAGAGCTTGGCGCGAATGGGCTACCGCTTCTCTGCTTTACGG 578  
 |||||  
 Qy 1121 TATCGCGCTCCCGATTTCGACGCGCATTCGCTTCTATCGCTTCTTGACGAGTCTT 1177  
 |||||  
 Db 579 TATCGCGCTCCCGATTTCGACGCGCATTCGCTTCTATCGCTTCTTGACGAGTCTT 635  
 |||||

RESULT 12  
 AQ398387  
 LOCUS

DEFINITION  
 mgxb0009J05f CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
 clone mgxb0009J05f, genomic survey sequence.  
 ACCESSION  
 AQ398387  
 VERSION  
 AQ398387.1 GI:4369414  
 KEYWORDS  
 GSS.  
 SOURCE  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Magnaporthe grisea

REFERENCE  
 1 (bases 1 to 620)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: TAATACGACTCCTATAGGG  
 Class: BAC ends  
 High quality sequence stop: 187.

JOURNAL  
 COMMENT

ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Magnaporthe grisea  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 620)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: TAATACGACTCCTATAGGG  
 Class: BAC ends  
 High quality sequence stop: 187.

```
FEATURES
  source      Location/Qualifiers
    1..620
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb0009J05f"
      /tissue_type="Protoplasts"
      /lab_host="E. coli DH10B"
      /clone_lib="CUGI Rice Blast BAC Library"
      /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
      Rice blast is one of the most devastating fungal diseases
      of rice world wide. It is a filamentous ascomycete with
      a haploid genome (n=7) of approximately 40 Mbp. Rice
      blast is an important model fungal pathogen for studying
      numerous aspects of the fungal-host interaction. In
      order to facilitate genome wide analysis, a BAC library
      containing 9216 clones with an average insert size of 130
      kbp was constructed. This library represents greater
      than 25X genome coverage. High density colony filters
      are available upon request."

ORIGIN
  Query Match      7.0%; Score 559.8; DB 8; Length 620;
  Best Local Similarity 98.8%; Pred. No. 3.2e-119;
  Matches 564; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 609 GTTCTTGGCAGCTGTGTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTG 668
DB 1 GTTCTTGGCAGCTGTGTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTG 60
QY 669 GCGAAGTCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCTGCTGCGGAGAACTATCC 728
DB 61 GCGAAGTCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCTGCTGCGGAGAACTATCC 120
QY 729 ATCATGGCTGATGCAATCGCGGCTGCTAGCTGTGATCCGGCTACCTGCCATTCGAC 788
DB 121 ATCATGGCTGATGCAATCGCGGCTGCTAGCTGTGATCCGGCTACCTGCCATTCGAC 180
QY 789 CACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGGCTGTGCGAT 848
DB 181 CACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGGCTGTGCGAT 240
QY 849 CAGGATGATCTGACGAAGCATCAGGGGCTCGCGCCAGCGAACTGTTCGCCAGGCTC 908
DB 241 CAGGATGATCTGACGAAGCATCAGGGGCTCGCGCCAGCGAACTGTTCGCCAGGCTC 300
QY 909 AAGGGCGCATGCCCGACGCGAGGATCTCGTCTGACCCATGGCGATGCTGTCGCG 968
DB 301 AAGGGCGCATGCCCGACGCGAGGATCTCGTCTGACCCATGGCGATGCTGTCGCG 360
QY 969 AATATCATGGTGAATAAGCGCTTTCTGGAATTCATGACTGTGGCGGCTGGGTGTG 1028
DB 361 AATATCATGGTGAATAAGCGCTTTCTGGAATTCATGACTGTGGCGGCTGGGTGTG 420
QY 1029 GCGGACCGCTATCAGACATAGCTTGTGCTACCGCTATTCCTGAAGGCTTGGCGCG 1088
DB 421 GCGGACCGCTATCAGACATAGCTTGTGCTACCGCTATTCCTGAAGGCTTGGCGCG 480
QY 1089 GAATGGGCTGACCGCTTCTCTGCTTTTACGGTATCGCGGCTCCCGATTCGACGCGCATC 1148
DB 481 GAATGGGCTGACCGCTTCTCTGCTTTTACGGTATCGCGGCTCCCGATTCGACGCGCATC 540
QY 1149 GCCTTCTATCGCTTCTTGACGAGTCTTCT 1179
DB 541 GCCTTCTATCGGCTTCTTGACGAGTCTTCTT 571

RESULT 13
CK801630
LOCUS
DEFINITION
  CK801630 549 bp mRNA linear EST 26-FEB-2004
  NF32c08f44.r1 Tall Rescue PI283316 44 deg C Heat Stress SSH cDNA
  Schedonorus arundinaceus cDNA clone NF32c08f44 5', mRNA sequence.

ACCESSION CK801630
VERSION CK801630.1 GI:43400943
KEYWORDS EST.
SOURCE Schedonorus arundinaceus (Festuca arundinacea)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Poae; Schedonorus.
REFERENCE 1 (bases 1 to 549)
AUTHORS Zhang,Y., Zwonitzer,J.C., Chekhovskiy,K., May,G.D. and Mian,M.A.R.
TITLE A functional genomics approach for identification of heat tolerance
genes in tall rescue
JOURNAL (in) Hopkins,A., Wang,Z.Y., Mian,R., Sledge,M. and Barker,R.E.
(edn.), MOLECULAR BREEDING OF FORAGE AND TURF, Kluwer Acad. Pub. 0,
87-96 (2003)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Contact Dr.Rouf Mian (rmian@noble.org) regarding clone availability
Seq primer: M13 reverse primer
High quality sequence stop: 470.

FEATURES
  source      Location/Qualifiers
    1..549
      /organism="Schedonorus arundinaceus"
      /mol_type="mRNA"
      /cultivar="PI283316"
      /db_xref="taxon:4606"
      /clone="NF32c08f44"
      /tissue_type="shoot"
      /lab_host="E.coli"
      /clone_lib="Tall Rescue PI283316 44 deg C Heat Stress SSH
      cDNA"
      /note="Vector: PCR 2.1-TOPO; Site 1: EcoR I; Site 2: EcoR
      I; Bp/Clontech PCR-select cDNA subtraction library"

ORIGIN
  Query Match      6.8%; Score 542.4; DB 7; Length 549;
  Best Local Similarity 99.8%; Pred. No. 3.6e-115;
  Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 556 TGAATGAACTGACGACGAGGAGGCGGCTATCGTGGCTGCGCAGCGGCTTCCTT 615
DB 6 TGAATGAACTGACGACGAGGAGGCGGCTATCGTGGCTGCGCAGCGGCTTCCTT 65
QY 616 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAAG 675
DB 66 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAAG 125
QY 676 TCGCGGGGCGAGATCTTCGTTCATCTACCTTGTCTCTCCGAGAAAGTATCATCATGG 735
DB 126 TCGCGGGGCGAGATCTTCGTTCATCTCCACCTTGTCTCTCCGAGAAAGTATCATCATGG 185
QY 736 CTGATGCATGCGGGGCTGCTATCGCTTGATCCGCTACCTGCGCATTCGACCCAAAG 795
DB 186 CTGATGCATGCGGGGCTGCTATCGCTTGATCCGCTACCTGCGCATTCGACCCAAAG 245
QY 796 CGAAACATCGCATCGAGCGAGCAGCTACTCTCGATGGAAGCCGCTCTTGTGATCAGGATG 855
DB 246 CGAAACATCGCATCGAGCGAGCAGCTACTCTCGATGGAAGCCGCTCTTGTGATCAGGATG 905
QY 856 ATCTGGAACGAAGCATCAGGGGCTGCGCCAGCGAACTGTTCGACGCTCAAGGCGC 915
DB 306 ATCTGGAACGAAGCATCAGGGGCTGCGCCAGCGAACTGTTCGACGCTCAAGGCGC 365
QY 916 GCATGCCCGACGCGAGGATCTCGCTGACCCATCGCGATGCTCTCTTGCAGATATCA 975
DB 366 GCATGCCCGACGCGAGGATCTCGCTGACCCATCGCGATGCTCTCTTGCAGATATCA 425
QY 976 TGGTGGAAATGCCCGCTTTTCTGGATTATCATGACTGTGCGCGGCTGGGTGTGGCGGACC 1035
```

```

Db      426  TGGTGAATAAGCGCGCTTTCTCGATTTCATCGACTGGCCGGCTGGGTGGGGACC 485
QY      1036  GCTATCAGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGCGAATGGG 1095
Db      486  GCTATCAGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGCGAATGGG 545
QY      1096  CTGA 1099
Db      546  CTGA 549

RESULT 14
LOCUS   AQ447874
DEFINITION  mgxb0012101f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
ACCESSION  AQ447874
VERSION    AQ447874.1 GI:4577011
KEYWORDS   GSS.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
1 (bases 1 to 561)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL    Unpublished (1998)
COMMENT     Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 42
High quality sequence stop: 326.
Location/Qualifiers
1. 561
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/db_xref="taxon:148305"
/clone="mgxb0012101f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
numerous aspects of the fungal pathogen for studying
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."

ORIGIN
Query Match 6.4%; Score 514.8; DB 8; Length 561;
Best Local Similarity 99.6%; Pred. No. 9.9e-109;
Matches 516; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      510  CAGGGGCGCCGGTCTTTTGTCAAGACCGAGCTGTCCGGTGCCTGAATGAATCGAG 569
Db      44   CAGGGGCGCCGGTCTTTTGTCAAGACCGAGCTGTCCGGTGCCTTTTGAATCGAG 103
QY      570  GACGAGGCGAGCGGCTATCGTGGCTGGCCAGCGGCGTCTTGGCGAGCTGTGCTC 629

```

```

Db      104  GACGAGGCGAGCGCGCTATCGTGGCTGGCCACGACGGGGGTTCTTGGCGAGCTGTGCTC 163
QY      630  GACGTTGTCACTGAAGCGGGAAGGAGCTGGTGTCTATTGGGCGAAAGTGCCTGGGGCAGGAT 689
Db      164  GAGGTTGTCACTGAAGCGGGAAGGAGCTGGTGTCTATTGGGCGAAAGTGCCTGGGGCAGGAT 223
QY      690  CTCCTGTCACTCACTTGTCTCTCCCGAGAAAGATATCCATCATCGCTGATGCAATGCGG 749
Db      224  CTCCTGTCACTCACTTGTCTCTCCCGAGAAAGATATCCATCATCGCTGATGCAATGCGG 283
QY      750  CGGCTGCATACGCTTGTATCCGCTACCTGACCCATTCGACCCACCAAGCGGAATCGCATC 809
Db      284  CGGCTGCATACGCTTGTATCCGCTACCTGACCCATTCGACCCACCAAGCGGAATCGCATC 343
QY      810  GAGCGAGCAGCTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAG 869
Db      344  GAGCGAGCAGCTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAG 403
QY      870  CATGAGGGGCTCGCGCCAGCCGAAGTGTTCGCGAGGCTCAAGGCGCGATGCCGACGGC 929
Db      404  CATGAGGGGCTCGCGCCAGCCGAAGTGTTCGCGAGGCTCAAGGCGCGATGCCGACGGC 463
QY      930  GAGGATCTCGTCTGACCCATCGGATGCGCTCTTGTGATCAGGATGATCTGGGAAATGCG 989
Db      464  GAGGATCTCGTCTGACCCATCGGATGCGCTCTTGTGATCAGGATGATCTGGGAAATGCG 523
QY      990  CGCTTTCTCGGATTCATCGACTGTGGCGCGCTGGGTGT 1027
Db      524  CGCTTTCTCGGATTCATCGACTGTGGCGCGCTGGGTGT 561

```

```

RESULT 15
LOCUS   AQ447775
DEFINITION  mgxb0011E13f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
ACCESSION  AQ447775
VERSION    AQ447775.1 GI:4576912
KEYWORDS   GSS.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
1 (bases 1 to 509)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL    Unpublished (1998)
COMMENT     Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 342.
Location/Qualifiers
1. 509
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/db_xref="taxon:148305"
/clone="mgxb0011E13f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with

```

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

ORIGIN

Query Match	6.4%;	Score 508;	DB 8;	Length 509;
Best Local Similarity	100.0%;	Pred. No. 3.8e-107;		
Matches 508;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

Qy	629	CGAGGTTGTACATGAGCGGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCGGGCAGGA	688
Db	1	CGAGGTTGTACATGAGCGGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCGGGCAGGA	60
Qy	689	TCTCCTGTCTATCTCACCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATGCG	748
Db	61	TCTCCTGTCTATCTCACCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATGCG	120
Qy	749	GGGCTGATACCGTTGATCCGGCTACTGCCATTCGACCAACAGCGAAACATCGCAT	808
Db	121	GGGCTGATACCGTTGATCCGGCTACTGCCATTCGACCAACAGCGAAACATCGCAT	180
Qy	809	CGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGACGAAGA	868
Db	181	CGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGACGAAGA	240
Qy	869	GCATCAGGGGCTCGCGCAGCCGAACTGTTCGCCAGGCTCAAGGCGGCGCATGCCGACGG	928
Db	241	GCATCAGGGGCTCGCGCAGCCGAACTGTTCGCCAGGCTCAAGGCGGCGCATGCCGACGG	300
Qy	929	CGAGGATCTCGTCCGTGACCCATGCGGATGCTGCTCCGAAATATCATGTGGAAATGG	988
Db	301	CGAGGATCTCGTCCGTGACCCATGCGGATGCTGCTCCGAAATATCATGTGGAAATGG	360
Qy	989	CCGCTTTTCTGGATTCTGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT	1048
Db	361	CCGCTTTTCTGGATTCTGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT	420
Qy	1049	AGCGTTGGCTACCGTGTATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT	1108
Db	421	AGCGTTGGCTACCGTGTATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT	480
Qy	1109	CGTCTTTTACGGTATCCCGCTCCCGAT	1136
Db	481	CGTCTTTTACGGTATCCCGCTCCCGAT	508

Search completed: December 9, 2004, 03:36:21  
Job time : 21856 secs